ETH Zürich
Priv.-Doz. Dr. Cornelia Busch

HS 2015

Mathematics

D-ARCH

Contents

1 Calculus **3**
1.1 Limits . . . 3
1.2 Differential calculus . . . 4
1.2.1 Definition . . . 4
1.2.2 Rules and examples . . . 5
1.2.3 Mean value theorem . . . 7
1.3 Integration . . . 7
1.3.1 Definition . . . 7
1.3.2 Main theorem of integration theory . . . 9
1.3.3 Partial integration . . . 10
1.3.4 Integration by substitution . . . 11
1.3.5 Partial fraction decomposition . . . 13

2 Multivariable Calculus: first part **14**
2.1 Multivariable functions . . . 14
2.2 Curves . . . 14
2.2.1 Parameterized curves . . . 14
2.2.2 The length of a curve . . . 15
2.2.3 The work of a force along a curve . . . 18
2.2.4 Solids of revolution . . . 19
2.3 Scalar fields . . . 22
2.3.1 Partial derivatives . . . 22
2.3.2 The gradient . . . 23
2.3.3 The total differential . . . 24
2.3.4 A chain rule for partial derivatives . . . 24

3 Linear algebra **25**
3.1 Functions . . . 25
3.2 Linear equations . . . 26
3.2.1 Introduction . . . 26
3.2.2 Systems of equations . . . 26
3.3 Matrices . . . 28
3.3.1 Vector spaces . . . 28
3.3.2 Linear dependency . . . 31
3.3.3 Addition and product of matrices . . . 33
3.3.4 The system of equations . . . 35
3.3.5 The transpose of a matrix . . . 36
3.3.6 Determinant . . . 37
3.4 Eigenvalues and eigenvectors . . . 40

3.4.1 Change of basis 40
3.4.2 Scalings 42
3.4.3 Eigenvalues and eigenvectors 43

4 Multivariable Calculus: second part **44**
4.0.4 Local extrema 44
4.1 Surfaces 45
4.1.1 Parameterized surfaces 45
4.1.2 Coordinates 46
4.1.3 The area of a surface 47

5 Differential equations **50**
5.1 Definition 50
5.2 First order differential equations 50
5.2.1 Separation of variables 51
5.3 Linear differential equations 54
5.3.1 Variation of constants 55
5.3.2 Linear differential equations of order n with constant coefficients 57

6 Appendix **61**
6.1 Another chain rule for partial derivatives 61
6.2 The flux through a surface 63
6.3 The orthogonal group 64
6.3.1 Linear groups 64
6.3.2 Isometries 64
6.3.3 Reflections and rotations 66

1 Calculus

1.1 Limits

Definition 1.1. A *function* f from a set A to a set B is a rule that defines for every $x \in A$ a $y = f(x) \in B$. We write

$$\begin{array}{rccl} f: & A & \to & B \\ & x & \mapsto & y = f(x)\,. \end{array}$$

We call

- $A = \mathrm{dom}(f)$ the *domain* of f,
- B the *codomain* or *range* of f.
- $\mathrm{im}(f) = \{y \in B \mid \exists x \in A \text{ with } f(x) = y\}$ the *image* of f
- $\mathrm{graph}(f) = \{(x, y) \in A \times B \mid y = f(x)\}$ the *graph* of f.

Definition 1.2. Let $f : \mathrm{dom}(f) \to \mathbb{R}$ be a real function, and $I \subseteq \mathrm{dom}(f)$ an open interval. Let $\overline{I}$ be the union of I with its boundaries. Let $\xi \in \overline{I}$. We say that the *limit of* $f(x)$ *as* x *approaches* ξ *is* η and write

$$\lim_{x \to \xi} f(x) = \eta$$

if the following statement is true. For every $\varepsilon > 0$, $\varepsilon \in \mathbb{R}$, there exists a $\delta > 0$, $\delta \in \mathbb{R}$, such that for every $x \in I$

$$\text{if } 0 < |x - \xi| < \delta \quad \text{then} \quad |f(x) - \eta| < \varepsilon\,.$$

If $\xi \in \mathrm{dom}(f)$, then the function f is *continuous* in ξ if and only if

$$\lim_{x \to \xi} f(x) = f(\xi)\,.$$

Example. We define

$$\begin{array}{rrcl} f: & (-\infty, 0) \cup (0, \infty) & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

with

$$f(x) = \begin{cases} -1 & \text{for } x < 0 \\ 1 & \text{for } x > 0\,. \end{cases}$$

Then for $x \in (-\infty, 0)$

$$\lim_{x \to 0} f(x) = -1$$

and for $x \in (0, \infty)$

$$\lim_{x \to 0} f(x) = 1\,.$$

This shows that there doesn't exist a continuous function $g : \mathbb{R} \to \mathbb{R}$ with $g(x) = f(x)$ for all $x \in \mathrm{dom}(f)$.

1.2 Differential calculus

1.2.1 Definition

Definition 1.3. Let

$$\begin{array}{rccc} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

be a function from $\mathbb{R}$ to $\mathbb{R}$. The *derivative* of f is defined to be the limit

$$\lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} =: f'(x)$$

Another notation for f' is

$$f'(x) = \frac{d}{dx} f(x)\,.$$

The derivative is not defined for each function or for every point $(x, f(x))$. Let $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto f(x) := |x|$ with $|x| = -x$ for $x \leq 0$ and $|x| = x$ for $0 \leq x$. Then the derivative is -1 for $x < 0$ and 1 for $0 < x$, but the derivative is not defined in $x = 0$. The derivative measures how the value of the function changes in the neighbourhood of x.

Remark 1.4. Let

$$\begin{array}{rccc} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

be a function that is differentiable in $x_0 \in \mathbb{R}$, i.e. the derivative

$$\frac{d}{dx} f(x_0)$$

exists. If

$$\frac{d}{dx} f(x_0) = f'(x_0) = 0$$

and the f' ist differentiable in x_0 then

$$\begin{cases} \frac{d^2}{dx^2} f(x_0) = f''(x_0) > 0 & \Rightarrow x_0 \text{ is a local minimum} \\ \frac{d^2}{dx^2} f(x_0) = f''(x_0) < 0 & \Rightarrow x_0 \text{ is a local maximum}\,. \end{cases}$$

No general statement is possible for $f''(x_0) = 0$. The function $f(x) = x^3$ has a saddle point in $x_0 = 0$ and $f(x) = x^4$ has a minimum in $x_0 = 0$.

1.2.2 Rules and examples

Some rules are well-known. Let $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$, $h : \mathbb{R} \to \mathbb{R}$ be functions and $a, b \in \mathbb{R}$. We assume that the derivatives of f, g and h are defined.
The derivative of a function f in x gives the slope of the tangent to the curve $\big(x, f(x)\big)$.

Properties		
sum	$(f+g)' = f' + g'$	
constant factor	$(\lambda f)' = \lambda f'$	
product	$(fg)' = f'g + fg'$	
quotient	$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$	
chain rule	$(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$	$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$
inverse	$g'(y) = \frac{1}{f'(g(y))}$	$\frac{dx}{dy} = \left(\frac{dy}{dx}\right)^{-1}$

Exponential functions, logarithm		
$f(x)$	$f'(x)$	Condition
c	0	c is a constant
x^n	nx^{n-1}	$n \in \mathbb{Z}$ and $x \neq 0$ if $n < 0$
x^a	ax^{a-1}	$a \in \mathbb{R}$ and $x > 0$
e^x	e^x	
a^x	$a^x \cdot \ln a$	$a > 0$
$\ln x$	$\frac{1}{x}$	$x > 0$

Trigonometric and hyperbolic functions			
$f(x)$	$f'(x)$	$f(x)$	$f'(x)$
$\sin x$	$\cos x$	$\sinh x$	$\cosh x$
$\cos x$	$-\sin x$	$\cosh x$	$\sinh x$
$\tan x$	$\frac{1}{\cos^2 x}$	$\tanh x$	$\frac{1}{\cosh^2 x}$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	$\operatorname{arsinh} x$	$\frac{1}{\sqrt{1+x^2}}$
$\arccos x$	$\frac{-1}{\sqrt{1-x^2}}$	$\operatorname{arcosh} x$	$\frac{1}{\sqrt{x^2-1}}$
$\arctan x$	$\frac{1}{1+x^2}$	$\operatorname{artanh} x$	$\frac{1}{1-x^2}$

Examples:

- $\frac{d}{dx}\frac{1}{\sin x}$

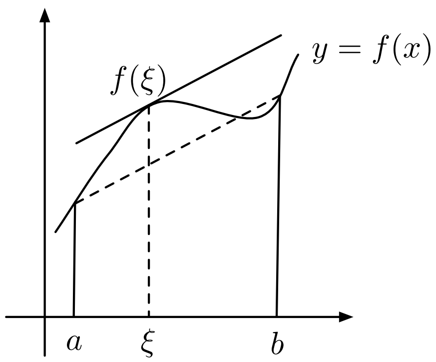
- $\frac{d}{dx} \sin \frac{1}{x}$

1.2.3 Mean value theorem

The mean value theorem of differential calculus is used for approximating functions.

Theorem 1.5. *Let $f : [a, b] \to \mathbb{R}$ be continuous and differentiable in the inner of $[a, b]$ (i.e. in (a, b)). Then there is a point $\xi \in (a, b)$ such that*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a} \quad \textit{resp.} \quad f(b) - f(a) = f'(\xi)(b - a)\,.$$



1.3 Integration

1.3.1 Definition

Given a function

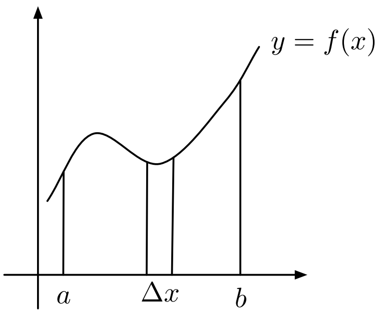
$$f:\ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & f(x) \end{array}$$

we compute the area between the graph $(x, f(x))$, $x \in [a, b]$ and the interval $[a, b]$, $a < b$.
We first cut the area in very thin stripes of width $\Delta x = \frac{b-a}{n}$ and approximate it with

$$\sum_{k=1}^{n} \Delta x\, f(\widetilde{x}_k)\,, \quad a + (k-1)\,\Delta x \le \widetilde{x}_k \le a + k\,\Delta x\,.$$

If

$$f(\widetilde{x}_k) = \max\{f(x) \mid a + (k-1)\,\Delta x \le x \le a + k\,\Delta x\}$$



we get an upper sum and if

$$f(\widetilde{x}_k) = \min\{f(x) \mid a + (k-1)\,\Delta x \le x \le a + k\,\Delta x\}$$

we get a lower sum.
We define the *integral of* f *on the interval* $[a, b]$

$$\int_a^b f(x)\,dx := \lim_{n\to\infty} \sum_{k=1}^{n} \Delta x\, f(\widetilde{x}_k)\,.$$

If the limit exists, then the limit of the upper sum equals the limit of the lower sum.

Definition 1.6. Let $f : \mathrm{dom}(f) \to \mathbb{R}$. If $[a, b] \subseteq \mathrm{dom}(f)$, then the integral

$$\int_a^b f(x)\,dx$$

is the *definite integral* of f on $[a, b]$.

We have

$$\int_{-\infty}^b f(x)\,dx = \lim_{a\to\infty} \int_{-a}^b f(x)\,dx$$

and

$$\int_a^{\infty} f(x)\,dx = \lim_{b\to\infty} \int_a^b f(x)\,dx$$

if the limits exist. Moreover

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_0^{\infty} f(x)\,dx\,.$$

If $\mathrm{dom}(f) = [a, b)$, then (with the assumption that the limits exist)

$$\int_a^b f(x)\,dx = \lim_{\varepsilon\to 0} \int_a^{b-\varepsilon} f(x)\,dx\,.$$

If $\mathrm{dom}(f) = (a, b]$, then (with the assumption that the limits exist)

$$\int_a^b f(x)\,dx = \lim_{\varepsilon\to 0}\int_{a+\varepsilon}^b f(x)\,dx\,.$$

Example.

$$\begin{aligned}\int_0^\infty \frac{1}{1+x^2}\,dx &= \lim_{b\to\infty}\int_0^b \frac{1}{1+x^2}\,dx = \lim_{b\to\infty}\left([\arctan x]_0^b\right)\\ &= \lim_{b\to\infty}\left(\arctan b - \arctan 0\right) = \pi/2\,.\end{aligned}$$

$$\begin{aligned}\int_0^\infty e^{-x}\,dx &= \lim_{b\to\infty}\int_0^b e^{-x}\,dx = \lim_{b\to\infty}\left(-e^{-b}+e^{-0}\right)\\ &= 1\,.\end{aligned}$$

1.3.2 Main theorem of integration theory

We define the area function

$$F_a : [a,\infty) \to \mathbb{R}$$

of a positive function $f : \mathbb{R} \to \mathbb{R}_0^+$ to be

$$F_a(x) = \int_a^x f(\tilde{x})d\tilde{x}\,,$$

the area between the graph of f and the x-axis on the interval $[a, x]$. This function satisfies $F_a(a) = 0$ and for $b > a$

$$F_b(x) = F_a(x) - F_a(b)$$

One can show that

$$F_a'(x) = f(x)\,.$$

Definition 1.7. A function $F : \mathrm{dom}(F) \to \mathbb{R}$, $\mathrm{dom}(F) \subseteq \mathbb{R}$ that satisfies

$$\frac{d}{dx}F(x) = F'(x) = f(x)$$

is called the *antiderivative* of f. Two different antiderivatives of f differ by a constant. The set of all antiderivatives of f is called the *indefinite integral* of f.

$$\begin{aligned}\int f(x)\,dx &:= \{F(x) \mid \tfrac{d}{dx}F(x) = f(x)\}\\ &= \{F(x) + c \mid c \in \mathbb{R},\ F \text{ is an antiderivative of } f\}\,.\end{aligned}$$

Theorem 1.8. *Given a function $f : \mathbb{R} \to \mathbb{R}$ with antiderivative $F : \mathbb{R} \to \mathbb{R}$. The definite integral of f on $[a, b]$ equals*

$$\int_a^b f(x)\,dx = F(b) - F(a)\,.$$

We also write

$$f(x)\,dx = dF$$

and

$$\int_a^b dF = F(b) - F(a)\,.$$

1.3.3 Partial integration

We know the product rule

$$\big(f(x) \cdot g(x)\big)' = f'(x) \cdot g(x) + f(x) \cdot g'(x)\,.$$

If we integrate

$$f'(x) \cdot g(x) = \big(f(x) \cdot g(x)\big)' - f(x) \cdot g'(x)$$

we get

$$\int f'(x) \cdot g(x)\,dx = f(x) \cdot g(x) - \int f(x) \cdot g'(x)\,dx\,.$$

This formula can be used to compute integrals of functions $f'(x) \cdot g(x)$.

Example. i) $\displaystyle\int x\,e^x\,dx = x\,e^x - \int e^x\,dx$

ii) $\displaystyle\int x^2\,e^x\,dx = x^2\,e^x - 2\int x\,e^x\,dx$

iii)

$$\int_0^{\pi/2} \cos x \cdot e^x\,dx = \cos x \cdot e^x\Big|_0^{\pi/2} + \int_0^{\pi/2} \sin x\,e^x\,dx$$
$$= \cos x \cdot e^x\Big|_0^{\pi/2} + \sin x \cdot e^x\Big|_0^{\pi/2} - \int_0^{\pi/2} \cos x \cdot e^x\,dx$$

iv) $\displaystyle\int \ln x\,dx = \int 1 \cdot \ln x\,dx$

1.3.4 Integration by substitution

This method uses the chain rule of differentiation. Let $x := \varphi(t)$. Then

$$\int f\big(\varphi(t)\big)\,\varphi'(t)\,dt = \left(\int f(x)\,dx\right)_{x:=\varphi(t)}$$

and

$$\int_a^b f\big(\varphi(t)\big)\,\varphi'(t)\,dt = \int_{\varphi(a)}^{\varphi(b)} f(x)\,dx\,.$$

Example. The integral

$$J := \int (\cos t + \cos^3 t)\,dt$$

can be transformed to

$$J = \int (1+\cos^2 t)\;\cos t\,dt = \int (2-\sin^2 t)\cos t\,dt\,.$$

The substitution

$$\sin t := x\,,\quad \cos t\,dt = dx$$

yields

$$\begin{aligned} J &= \left(\int (2-x^2)\,dx\right)_{x:=\sin t} = \left(2x - \frac{x^3}{3}\right)_{x:=\sin t} + c \\ &= 2\sin t - \frac{1}{3}\sin^3 t + c\,. \end{aligned}$$

The definite integral

$$J_0 := \int_{\pi/6}^{\pi} (\cos t + \cos^3 t)\,dt$$

becomes with this substitution with $\sin \pi/6 = 1/2$ and $\sin\pi = 0$

$$\begin{aligned} J_0 &= \int_{1/2}^{0} (2-x^2)\,dx = \left(2x - \frac{x^3}{3}\right)\Bigg|_{1/2}^{0} \\ &= -\left(1 - \frac{1}{24}\right) = -\frac{23}{24}\,. \end{aligned}$$

Example. We compute the integral

$$\int \frac{3x-1}{x^2-x+1}\,dx\,.$$

We know that $(x^2-x+1)' = 2x-1$ and since

$$\begin{aligned}\frac{3x-1}{x^2-x+1} &= \frac{3}{2}\cdot\frac{2x-\frac{2}{3}}{x^2-x+1} = \frac{3}{2}\cdot\frac{2x-1+\frac{1}{3}}{x^2-x+1}\\ &= \frac{3}{2}\Big(\frac{2x-1}{x^2-x+1}+\frac{\frac{1}{3}}{x^2-x+1}\Big)\end{aligned}$$

we get

$$\begin{aligned}\int \frac{3x-1}{x^2-x+1} &= \frac{3}{2}\int\Big(\frac{2x-1}{x^2-x+1}+\frac{\frac{1}{3}}{x^2-x+1}\Big)\\ &= \frac{3}{2}\ln|x^2-x+1| + \frac{1}{2}\int\frac{1}{x^2-x+1}\,dx\end{aligned}$$

We have to compute the second integral. With

$$x^2-x+1 = \Big(x-\frac{1}{2}\Big)^2+\frac{3}{4} = \frac{3}{4}\left(\left(\frac{x-\frac{1}{2}}{\sqrt{\frac{3}{4}}}\right)^2+1\right)$$

we get

$$\int\frac{1}{x^2-x+1}\,dx = \frac{4}{3}\int\frac{1}{\left(\frac{x-\frac{1}{2}}{\sqrt{\frac{3}{4}}}\right)^2+1}\,dx$$

and substitute

$$u = \frac{2}{\sqrt{3}}\Big(x-\frac{1}{2}\Big)\,,\quad du = \frac{2}{\sqrt{3}}\,dx\,.$$

Herewith

$$\frac{4}{3}\,\frac{\sqrt{3}}{2}\int\frac{1}{u^2+1}\,du = \frac{2}{\sqrt{3}}\arctan u + C$$

and

$$\int\frac{1}{x^2-x+1}\,dx = \frac{2}{\sqrt{3}}\arctan\Big(\frac{2}{\sqrt{3}}\Big(x-\frac{1}{2}\Big)\Big)+C\,.$$

Therefore

$$\int\frac{3x-1}{x^2-x+1} = \frac{3}{2}\ln|x^2-x+1| + \frac{1}{\sqrt{3}}\arctan\Big(\frac{2}{\sqrt{3}}\Big(x-\frac{1}{2}\Big)\Big)+C\,.$$

Another method is the following. If the function $\varphi(t)$ is invertible on the required t-interval, then

$$\int f(x)\,dx = \left(\int f\big(\varphi(t)\big)\,\varphi'(t)\,dt\right)_{t:=\varphi^{-1}(x)}$$

and

$$\int_a^b f(x)\,dx = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f\big(\varphi(t)\big)\,\varphi'(t)\,dt\,.$$

We will see examples of this method in multivariable analysis.

1.3.5 Partial fraction decomposition

We consider functions

$$f(x) = \frac{P_n(x)}{Q_m(x)}\,,$$

where $P_n(x)$, $Q_m(x)$ are polynomials of degree $n < m$. Let $\alpha_1, \dots, \alpha_m$ be the zeros of $Q_m(x)$ and $P_n(\alpha_i) \neq 0$ for $i = 1, \dots, m$. If $\alpha_1, \dots, \alpha_m$ are distinct constants, then we make the ansatz

$$\frac{P_n(x)}{Q_m(x)} = \frac{A_1}{x-\alpha_1} + \frac{A_2}{x-\alpha_2} + \dots + \frac{A_m}{x-\alpha_m}\,.$$

If the multiplicity of α_2 is $\ell > 1$, then we make the ansatz

$$\frac{P_n(x)}{Q_m(x)} = \frac{A_1}{x-\alpha_1} + \frac{B_1}{x-\alpha_2} + \frac{B_2}{(x-\alpha_2)^2} + \dots + \frac{B_\ell}{(x-\alpha_2)^\ell} + \frac{C_1}{x-\alpha_3} + \dots$$

Example.

- $\dfrac{-1}{(x+2)(x+3)} = \dfrac{A}{x+2} + \dfrac{B}{x+3}$

  with $A = -1$, $B = 1$.

- $\dfrac{x}{(x-1)^2\,(x+1)} = \dfrac{A_1}{x-1} + \dfrac{A_2}{(x-1)^2} + \dfrac{B}{x+1}$

  with $A_1 = \dfrac{1}{4}$, $A_2 = \dfrac{1}{2}$, $B = \dfrac{-1}{4}$.

2 Multivariable Calculus: first part

2.1 Multivariable functions

The functions studied in multivariable calculus are

Curves	$f : \mathbb{R} \to \mathbb{R}^n$	Length of curves, line integrals, curvature
Surfaces	$f : \mathbb{R}^2 \to \mathbb{R}^n$	Areas of surfaces, surface integrals, flux through surfaces, curvature
Scalar fields	$f : \mathbb{R}^n \to \mathbb{R}$	Maxima and minima, Lagrange multipliers, directional derivatives
Vector fields	$f : \mathbb{R}^m \to \mathbb{R}^n$	Any of the operations of vector calculus, gradient, divergence, curl

Example. The curve given by

$$\begin{array}{rccl} f: & \mathbb{R} & \longrightarrow & \mathbb{R}^2 \\ & t & \longmapsto & (\cos t, \sin t) \end{array}$$

is the unit circle in the plane.
The curve parameterized with

$$\begin{array}{rccl} f: & \mathbb{R} & \longrightarrow & \mathbb{R}^3 \\ & t & \longmapsto & (\cos t, \sin t, t) \end{array}$$

is a line that "screws up".

2.2 Curves

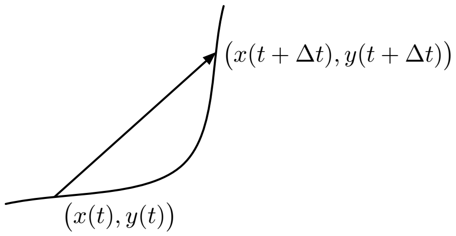
2.2.1 Parameterized curves

We already know vector-valued functions. Let I be an interval.

$$\begin{array}{rccl} \vec{r}: & I & \longrightarrow & \mathbb{R}^{3(2)} \\ & t & \longmapsto & \vec{r}(t) = (x(t), y(t), z(t)) \end{array}$$

Then $\vec{r}(t)$ is a point on a curve. In order to define the derivative we first consider the quotients:

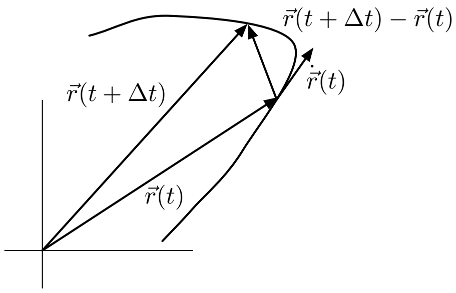
$$\begin{aligned} &\frac{1}{\Delta t}\big(\vec{r}(t+\Delta t) - \vec{r}(t)\big) \\ &\quad = \frac{1}{\Delta t}\Big(x(t+\Delta t) - x(t),\ y(t+\Delta t) - y(t),\ z(t+\Delta t) - z(t)\Big) \\ &\quad = \left(\frac{x(t+\Delta t) - x(t)}{\Delta t}, \frac{y(t+\Delta t) - y(t)}{\Delta t}, \frac{z(t+\Delta t) - z(t)}{\Delta t}\right) \end{aligned}$$



If the functions $x(t)$, $y(t)$ and $z(t)$ are differentiable, then the derivative of $\vec{r}$ is defined:

$$\frac{d\vec{r}}{dt}(t) := \lim_{\Delta t \to 0} \frac{\vec{r}(t+\Delta t) - \vec{r}(t)}{\Delta t} = \Big(\frac{dx}{dt}(t)\,,\ \frac{dy}{dt}(t)\,,\ \frac{dz}{dt}(t)\Big)\,.$$

Another notation is (used for example in physics) $\dot{\vec{r}}(t)$ for $\frac{d\vec{r}}{dt}(t)$.



The limit is the tangent vector $\frac{d}{dt}\vec{r}(t) = \dot{\vec{r}}(t)$ at the curve $\vec{r}$ in $r(t)$.

2.2.2 The length of a curve

Definition 2.1. Let

$$\begin{array}{cccc} \vec{r}\,: & [a,b] & \longrightarrow & \mathbb{R}^{3(2)} \\ & t & \longmapsto & \vec{r}(t) := (x(t), y(t), z(t)) \end{array}$$

be the parametric representation of a curve γ that joins $\vec{r}(a)$ with $\vec{r}(b)$. Let $\frac{d}{dt}\vec{r}(t) = \dot{\vec{r}}(t)$ be the derivative of $\vec{r}(t)$.
The *length* of the curve γ is

$$L(\gamma) := \int_{[a,b]} \left|\dot{\vec{r}}(t)\right| dt = \int_{[a,b]} ds$$

with

$$ds := \left|\dot{\vec{r}}(t)\right| dt = \sqrt{\dot{x}^2(t) + \dot{y}^2(t) + \dot{z}^2(t)}\, dt\,.$$

Example. We consider

$$\begin{array}{rccl} \vec{r}: & \mathbb{R} & \longrightarrow & \mathbb{R}^3 \\ & t & \longmapsto & \vec{r}(t) := (x(t), y(t), z(t)) = (\cos(t), \sin(t), t) \end{array}$$

Then the length of the curve between $\vec{r}(0)$ and $\vec{r}(2\pi)$ is

$$\begin{aligned} L(\gamma) &= \int_0^{2\pi} \left|\dot{\vec{r}}(t)\right| dt \\ &= \int_0^{2\pi} \left|\Big(\frac{d}{dt}\cos(t), \frac{d}{dt}\sin(t), \frac{d}{dt}\,t\Big)\right| dt \\ &= \int_0^{2\pi} \left|\big(-\sin(t), \cos(t), 1\big)\right| dt \\ &= \int_0^{2\pi} \sqrt{\sin^2(t) + \cos^2(t) + 1}\,dt \\ &= \int_0^{2\pi} \sqrt{2}\,dt = \sqrt{2}\,t\Big|_0^{2\pi} \\ &= 2\sqrt{2}\,\pi \end{aligned}$$

Example. We consider the curve

$$\begin{array}{rccl} \vec{r}: & \mathbb{R} & \longrightarrow & \mathbb{R}^2 \\ & t & \longmapsto & \vec{r}(t) := \big(x(t), y(t)\big) = \big(t\cos(t), t\sin(t)\big) \end{array}$$

The length of the curve between $\vec{r}(0)$ and $\vec{t}(2\pi)$ is

$$\begin{aligned} L(\gamma) &= \int_0^{2\pi} \left|\dot{\vec{r}}(t)\right| dt \\ &= \int_0^{2\pi} \left|\Big(\frac{d}{dt}\,t\cos(t), \frac{d}{dt}\,t\sin(t)\Big)\right| dt \\ &= \int_0^{2\pi} \left|\big(\cos(t) - t\sin(t),\ \sin(t) + t\cos(t)\big)\right| dt \\ &= \int_0^{2\pi} \sqrt{\big(\cos(t) - t\sin(t)\big)^2 + \big(\sin(t) + t\cos(t)\big)^2}\ dt \\ &= \int_0^{2\pi} \sqrt{\big(1+t^2\big)\big(\cos^2(t) + \sin^2(t)\big)}\,dt \\ &= \int_0^{2\pi} \sqrt{\big(1+t^2\big)}\,dt\,. \end{aligned}$$

With the substitution $t = \sinh s$ we get

$$
\begin{aligned}
L(\gamma) &= \int_{\operatorname{arsinh}(0)}^{\operatorname{arsinh}(2\pi)} \sqrt{1+\sinh^2(s)}\,\cosh(s)\,ds \\
&= \int_{\operatorname{arsinh}(0)}^{\operatorname{arsinh}(2\pi)} \cosh^2(s)\,ds = \int_{\operatorname{arsinh}(0)}^{\operatorname{arsinh}(2\pi)} \Big(\frac{e^s+e^{-s}}{2}\Big)^2\,ds \\
&= \frac{1}{4}\int_{\operatorname{arsinh}(0)}^{\operatorname{arsinh}(2\pi)} e^{2s}+e^{-2s}+2\,ds \\
&= \frac{1}{4}\Big(\frac{1}{2}e^{2s}-\frac{1}{2}e^{-2s}+2s\Big)\Big|_{\operatorname{arsinh}(0)}^{\operatorname{arsinh}(2\pi)} \\
&= \frac{1}{4}\Big(\frac{1}{2}e^{2s}-\frac{1}{2}e^{-2s}+2s\Big)\Big|_{0}^{\ln(2\pi+\sqrt{4\pi^2+1})} \\
&= \frac{1}{8}(2\pi+\sqrt{4\pi^2+1})^2-\frac{1}{8}-\frac{1}{8}(2\pi+\sqrt{4\pi^2+1})^{-2}+\frac{1}{8} \\
&\qquad +\frac{1}{2}\ln\big(2\pi+\sqrt{4\pi^2+1}\big) \\
&= \frac{1}{8}\left(\big(2\pi+\sqrt{4\pi^2+1}\big)^2-\frac{1}{\big(2\pi+\sqrt{4\pi^2+1}\big)^2}\right) \\
&\qquad +\frac{1}{2}\ln\big(2\pi+\sqrt{4\pi^2+1}\big) \\
&= \frac{1}{8}\left(\big(2\pi+\sqrt{4\pi^2+1}\big)^2-\frac{\big(2\pi-\sqrt{4\pi^2+1}\big)^2}{\big(4\pi^2-(4\pi^2+1)\big)^2}\right) \\
&\qquad +\frac{1}{2}\ln\big(2\pi+\sqrt{4\pi^2+1}\big) \\
&= \frac{1}{8}\Big(\big(4\pi^2+4\pi\sqrt{4\pi^2+1}+4\pi^2+1\big)-\big(4\pi^2-4\pi\sqrt{4\pi^2+1}+4\pi^2+1\big)\Big) \\
&\qquad +\frac{1}{2}\ln\big(2\pi+\sqrt{4\pi^2+1}\big) \\
&= \frac{1}{8}\big(8\pi\sqrt{4\pi^2+1}\big)+\frac{1}{2}\ln\big(2\pi+\sqrt{4\pi^2+1}\big) \\
&= \pi\sqrt{4\pi^2+1}+\frac{1}{2}\ln\big(2\pi+\sqrt{4\pi^2+1}\big)\,.
\end{aligned}
$$

Here we used

$$\operatorname{arsinh}(s) = \ln(s+\sqrt{s^2+1})\,.$$

2.2.3 The work of a force along a curve

Let

$$\begin{array}{rccl} \vec{F}: & \mathbb{R}^{3(2)} & \longrightarrow & \mathbb{R}^{3(2)} \\ & (x,y,z) & \longmapsto & \vec{F}(x,y,z) = \big(F_1(x,y,z), F_2(x,y,z), F_3(x,y,z)\big) \end{array}$$

be a force in every point of $\mathbb{R}^{3(2)}$, i.e. graviation.
The work of the force F along the curve γ parameterized with

$$\begin{array}{rccl} \vec{r}: & [a,b] & \longrightarrow & \mathbb{R}^{3(2)} \\ & t & \longmapsto & \vec{r}(t) = (x(t), y(t), z(t)) \end{array}$$

is defined to be the integral

$$W = \int_\gamma \vec{F} \cdot d\vec{r}\,,$$

where $\vec{F} \cdot d\vec{r}$ denotes the scalar product of $\vec{F}$ and the line element $d\vec{r}$. With the parameterization it can be written as

$$W = \int_a^b \vec{F}(\vec{r}(t)) \cdot \dot{\vec{r}}(t) dt\,.$$

Herewith

$$\begin{aligned} W &= \int_\gamma \vec{F} \cdot d\vec{r} \\ &= \int_a^b \vec{F}(\vec{r}(t)) \cdot \dot{\vec{r}}(t) dt \\ &= \int_a^b \big(F_1(\vec{r}(t)), F_2(\vec{r}(t)), F_3(\vec{r}(t))\big) \cdot \dot{\vec{r}}(t) dt \\ &= \int_a^b \big(F_1(\vec{r}(t)), F_2(\vec{r}(t)), F_3(\vec{r}(t))\big) \cdot \big(\dot{x}(t), \dot{y}(t), \dot{z}(t)\big) dt \\ &= \int_a^b F_1(\vec{r}(t))\dot{x}(t) + F_2(\vec{r}(t))\dot{y}(t) + F_3(\vec{r}(t))\dot{z}(t)\, dt \end{aligned}$$

Example. We consider the curve

$$\begin{array}{rccl} \vec{r}: & \mathbb{R} & \longrightarrow & \mathbb{R}^3 \\ & t & \longmapsto & \vec{r}(t) := (x(t), y(t), z(t)) = (\cos(t), \sin(t), t) \end{array}$$

and the force

$$\begin{array}{rccl} \vec{F}: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x,y,z) & \longmapsto & \vec{F}(x,y,z) = (-y, x, 0) \end{array}$$

The work of the force F along the curve $\gamma := \vec{r}(t)|_{[0,2\pi]}$ equals

$$\begin{aligned} W &= \int_\gamma \vec{F}\cdot d\vec{r} \\ &= \int_0^{2\pi} \vec{F}(\vec{r}(t))\cdot\dot{\vec{r}}(t)dt \\ &= \int_0^{2\pi} -y(t)\dot{x}(t)+x(t)\dot{y}(t)+0\cdot\dot{z}(t)\,dt \\ &= \int_0^{2\pi} (\sin^2(t)+\cos^2(t)+0)\,dt \\ &= \int_0^{2\pi} 1\,dt = t\Big|_0^{2\pi} \\ &= 2\pi\,. \end{aligned}$$

2.2.4 Solids of revolution

Let $[a,b]\subset\mathbb{R}$ and

$$f:[a,b]\longrightarrow\mathbb{R}\,,\qquad x\longmapsto f(x)=:y$$

be a real function and

$$\{(x,f(x))=(x,y)\mid x\in[a,b]\}$$

the graph of f. We consider the body obtained by revolution of the graph of f about the x-axis. In order to compute the volume of the body we cut discs that are Δx (dx) thick and have the radius $f(x)$. Then we get the volume

$$V_x=\pi\int_a^b (f(x))^2\,dx\,.$$

If the curve is given by a parametric representation, then

$$V_x=\pi\int_{t_1}^{t_2} y^2(t)\dot{x}(t)\,dt\,.$$

If the function f is invertible, i.e. if for every

$$y\in[\min\{f(a),f(b)\},\max\{f(a),f(b)\}]$$

there is a unique $x\in[a,b]$ with $y=f(x)$, then the volume of the solid of revolution around the y-axis is

$$V_y=\pi\int_{\min\{f(a),f(b)]\}}^{\max\{f(a),f(b)\}} (f^{-1}(y))^2\,dy\,.$$

With the substitution $x = f^{-1}(y)$ we get

$$V_y = \pi \int_{\min\{f(a),f(b)]\}}^{\max\{f(a),f(b)\}} x^2\,dy = \pi \int_a^b x^2 \cdot |f'(x)|\,dx\,.$$

In order to compute the surface of the body we consider stripes that are ds wide and $2\pi y$ long. Then

$$dS_x = 2\pi y ds$$

for the revolution about the x-axis and

$$dS_y = 2\pi x ds$$

for the rotation about the y-axis. We get the formulas

$$S_x = 2\pi \int_\alpha^\beta y(t)\sqrt{\dot{x}^2 + \dot{y}^2}\,dt$$

for the rotation about the x-axis and

$$S_y = 2\pi \int_\alpha^\beta x(t)\sqrt{\dot{x}^2 + \dot{y}^2}\,dt$$

for the rotation about the y-axis.

Example. To obtain the ellipsoid of revolution, we let the ellipse

$$x = a\cos t, \quad y = b\sin t$$

rotate about the x-axis. Then

$$\begin{aligned}
\frac{1}{2}V_x &= \pi \int_{\frac{\pi}{2}}^0 y^2(t)\dot{x}(t)\,dt \\
&= \pi \int_{\frac{\pi}{2}}^0 b^2 \sin^2 t(-a\sin t)\,dt \\
&= \pi ab^2 \int_0^{\frac{\pi}{2}} \sin^2 t\, \sin t\,dt \\
&= \pi ab^2 \int_0^{\frac{\pi}{2}} (1-\cos^2 t)\, \sin t\,dt \\
&= \pi ab^2 \int_0^{\frac{\pi}{2}} \sin t\,dt + \pi ab^2 \int_0^{\frac{\pi}{2}} -\cos^2 t\, \sin t\,dt
\end{aligned}$$

Substituting $s = \cos t$ in the second integral we get

$$\begin{aligned}\frac{1}{2}V_x &= -\pi ab^2 \cos t\Big|_0^{\frac{\pi}{2}} + \pi ab^2 \int_1^0 s^2 ds \\ &= -\pi ab^2(0-1) + \pi ab^2\,\frac{1}{3}(0-1) \\ &= \frac{2}{3}\pi ab^2\end{aligned}$$

and the volume is

$$V_x = \frac{4\pi}{3}ab^2\,.$$

Rotation around the y-axis yields

$$V_y = \frac{4\pi}{3}a^2 b\,.$$

Example. The surface of a sphere can be generated by revolving the upper half of the circle $x^2 + y^2 = r^2$ about the x-axis. Here $y = \sqrt{r^2 - x^2}$. We have $dS = 2\pi y ds$. The surface area of the sphere is

$$\begin{aligned}S &= \int_{-r}^{r} 2\pi y\, ds = \int_{-r}^{r} 2\pi y \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \\ &= 2\pi \int_{-r}^{r} y \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx\,.\end{aligned}$$

Differentiating the equation $x^2 + y^2 = r^2$ we get

$$2x + 2y \cdot \frac{dy}{dx} = 0 \quad \text{or} \quad \frac{dy}{dx} = -\,\frac{x}{y}\,.$$

We get

$$\begin{aligned}S &= 2\pi \int_{-r}^{r} y \sqrt{1 + \left(\frac{x^2}{y^2}\right)}\, dx = 2\pi \int_{-r}^{r} y \sqrt{\frac{x^2 + y^2}{y^2}}\, dx \\ &= 2\pi \int_{-r}^{r} y \sqrt{\frac{r^2}{y^2}}\, dx = 2\pi \int_{-r}^{r} r\, dx\,,\end{aligned}$$

where r is a constant and therefore

$$\begin{aligned}S &= 2\pi r \int_{-r}^{r} dx = 2\pi r x|_{-r}^{r} = 2\pi r(r - (-r)) \\ &= 4\pi r^2\,.\end{aligned}$$

2.3 Scalar fields

In this section we consider functions

$$\begin{array}{rccc} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

that map points $x = (x_1, \dots, x_n)$ in $\mathbb{R}^n$ to scalars $f(x_1, \dots, x_n)$. If $D \subset \mathbb{R}^n$, then the graph

$$\big\{\big(x, f(x)\big) \in \mathbb{R}^n \times \mathbb{R} \mid x \in D\big\}$$

describes a surface over D.
The function f may represent the metres above sea level of a point on a map or the temperature at a point in a space.

Definition 2.2. The *level set* of the function

$$\begin{array}{rccc} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

to the level c, $c \in \mathbb{R}$, is the set

$$f^{-1}(c) := \big\{(x_1, \dots, x_n) \in \mathbb{R}^n \mid f(x_1, \dots, x_n) = c\big\},$$

On the examples above it corresponds to the points at the same altitude or with the same temperature.

Example. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto f(x, y) = e^{-(x^2+y^2)}$. What are its level sets?

2.3.1 Partial derivatives

We fix a point $x^0 = (x_1^0, \dots, x_n^0) \in \mathbb{R}^n$ and consider the line

$$L_i := \{(x_1^0, \dots, x_{i-1}^0, x_i, x_{i+1}^0, \dots, x_n^0) \in \mathbb{R}^n \mid x_i \in \mathbb{R}\}.$$

Then the set

$$\mathcal{C}_i(x^0) := \{(x_1^0, \dots, x_{i-1}^0, x_i, x_{i+1}^0, \dots, x_n^0, f(\dots x_{i-1}^0, x_i, x_{i+1}^0, \dots)) \mid x_i \in \mathbb{R}\}$$

is a curve over the line L_i. It is the graph of the function

$$\begin{array}{rccc} \varphi_i & : \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x_i & \longmapsto & f(x_1^0, \dots, x_{i-1}^0, x_i, x_{i+1}^0, \dots, x_n^0). \end{array}$$

We consider the derivative of φ_i with respect to the variable x_i in the point x^0. This is called the *partial derivative* of f in x^0 with respect to x_i and written

$$f_i(x^0) \qquad \text{or} \qquad \frac{\partial f}{\partial x_i}(x^0)\,.$$

It is defined to be the limit

$$f_i(x^0) := \lim_{\Delta x \to 0} \frac{f(x_1^0, \ldots, x_i^0 + \Delta x, \ldots, x_n^0) - f(x_1^0, \ldots, x_i^0, \ldots, x_n^0)}{\Delta x}\,.$$

The tangent at $\mathcal{C}_i(x^0)$ in $p = \big(x^0, f(x^0)\big)$ is given by

$$\left\{ \big(x^0, f(x^0)\big) + \frac{\partial}{\partial x_i} f(x^0)(x_i - x_i^0) \,\middle|\, x_i \in \mathbb{R} \right\}.$$

The tangents $\mathcal{C}_i(x^0)$, $i = 1, \ldots, n$, span the tangent vector space

$$T_p S$$

at the surface S in p.

Example. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto f(x, y) = e^{-(x^2+y^2)}$. Determine the partial derivatives.

2.3.2 The gradient

Definition 2.3. We assume that all partial derivatives $\frac{\partial}{\partial x_i} f$, $i = 1, \ldots, n$ of the function

$$\begin{array}{rccc} f: & D & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

exist and that they are continuous. Then the vector

$$\nabla f(x^0) := \left(\frac{\partial}{\partial x_1} f(x^0), \ldots, \frac{\partial}{\partial x_n} f(x^0) \right)$$

is defined and called the *gradient* of f in x^0.

Theorem 2.4. *The gradient $\nabla f(x^0)$ is perpendicular to the level set*

$$f^{-1}\big(f(x^0)\big) := \big\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid f(x_1, \ldots, x_n) = f(x^0)\big\}\,.$$

Remark 2.5. The gradient ∇f is also denoted $\mathrm{grad}(f)$. It indicates the direction of maximal slope.

2.3.3 The total differential

Definition 2.6. The *total differential* of a function $f : \mathbb{R}^3 \to \mathbb{R}$ in 3 variables in (x_0, y_0, z_0) is defined to be

$$df = f_x(x_0, y_0, z_0)\, dx + f_y(x_0, y_0, z_0)\, dy + f_z(x_0, y_0, z_0)\, dz\,.$$

The total differential of a function in 2 variables in (x_0, y_0) is

$$df = f_x(x_0, y_0)\, dx + f_y(x_0, y_0)\, dy\,.$$

2.3.4 A chain rule for partial derivatives

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a differentiable function. Consider the composition

$$f\big(x(s,t), y(s,t)\big)$$

for differentiable functions $x, y : \mathbb{R}^2 \to \mathbb{R}$. Then

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial s}$$
$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial t}$$

We show the first equation. We consider the limit

$$\begin{aligned}
&\lim_{\Delta s \to 0} \frac{1}{\Delta s}\Big(f\big(x(s+\Delta s, t), y(s+\Delta s, t)\big) - f\big(x(s,t), y(s,t)\big)\Big) \\
&\quad = \lim_{\Delta s \to 0} \frac{1}{\Delta s}\Big(f\big(x(s+\Delta s, t), y(s+\Delta s, t)\big) - f\big(x(s,t), y(s+\Delta s,t)\big)\Big) \\
&\qquad + \lim_{\Delta s \to 0} \frac{1}{\Delta s}\Big(f\big(x(s,t), y(s+\Delta s, t)\big) - f\big(x(s,t), y(s,t)\big)\Big) \\
&\quad \approx \lim_{dx \to 0} \frac{\frac{\partial x}{\partial s}}{dx}\Big(f\big(x(s,t)+dx, y(s+\Delta s, t)\big) - f\big(x(s,t), y(s+\Delta s,t)\big)\Big) \\
&\qquad + \lim_{dy \to 0} \frac{\frac{\partial y}{\partial s}}{dy}\Big(f\big(x(s,t), y(s,t)+dy\big) - f\big(x(s,t), y(s,t)\big)\Big) \\
&\quad = \frac{\partial f}{\partial x}\big(x(s,t), y(s,t)\big) \cdot \frac{\partial x}{\partial s}(s,t) + \frac{\partial f}{\partial y}\big(x(s,t), y(s,t)\big) \cdot \frac{\partial y}{\partial s}(s,t)
\end{aligned}$$

The proof of the rule for the partial derivative with respect to t is analogous. For the special case $x(t)$, $y(t)$ we have

$$\frac{df}{dt}\big(x(t), y(t)\big) = \frac{\partial f}{\partial x}\big(x(t), y(t)\big) \cdot x'(t) + \frac{\partial f}{\partial y}\big(x(t), y(t)\big) \cdot y'(t)\,.$$

3 Linear algebra

3.1 Functions

Linear algebra is the theory of linear functions.

Definition 3.1. Let

$$\begin{array}{rccc} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) \end{array}$$

be a function mapping the real numbers on real numbers. The function f is called *linear* if and only if for all $x, y \in \mathbb{R}$ and $\lambda \in \mathbb{R}$

$$f(\lambda x + y) = \lambda f(x) + f(y) .$$

The definition for functions $\mathbb{Q} \to \mathbb{Q}$ on the rational numbers or functions $\mathbb{C} \to \mathbb{C}$ on the complex numbers is analogous.

Example. Let $a \in \mathbb{R}$ (or $a \in \mathbb{Q}$, $a \in \mathbb{C}$) be a constant. Then the function given by

$$f(x) = ax, \qquad x \in \mathbb{R}\,(\mathbb{Q}, \mathbb{C})$$

is a linear function.

Definition 3.2. Let

$$\begin{array}{rccc} g: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & x & \longmapsto & g(x) \end{array}$$

be a function mapping a vector with n components

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

on vectors with m components (coordinates)

$$g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix} = \begin{pmatrix} g_1(x_1, \ldots, x_n) \\ \vdots \\ g_m(x_1, \ldots, x_n) \end{pmatrix} ,$$

where

$$\begin{array}{rccc} g_i: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & g_i(x) , \end{array}$$

$i = 1, \ldots, m$, is a function in m variables. The sets $\mathbb{R}^n$ and $\mathbb{R}^m$ are *vector spaces*. The natural number n is called the *dimension* of the vector space $\mathbb{R}^n$. The function g is called *linear* if and only if for all $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$

$$g(\lambda x + y) = \lambda g(x) + g(y) .$$

The function g is linear if and only if the funtions g_i, $i = 1, \ldots, m$, are linear.

Example. The function

$$\begin{array}{rccc} ||\cdot||: & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & ||x|| := \sqrt{x_1^2 + \cdots + x_n^2}\,, \end{array}$$

is the norm of the vector x. The norm is not linear.

3.2 Linear equations

3.2.1 Introduction

Given $a, b \in \mathbb{R}$, ($\mathbb{Q}$, or $\mathbb{C}$). When does the linear equation in x

$$a \cdot x = b$$

have a solution? Is there more than one solution?
Let a and b be rational, real or complex numbers. If it is possible to divide by a, i.e. $a \neq 0$ or a is invertible, then the equation has a unique solution

$$x = a^{-1} \cdot b = \frac{b}{a}\,.$$

If $a = 0$ and $b = 0$ we have the equation

$$0 \cdot x = 0$$

and this is true for any value of x. Therefore the equation $0 \cdot x = 0$ has infinitely many solutions. If $a = 0$ and $b \neq 0$, then the equation

$$0 \cdot x = b \neq 0$$

has no solutions.
Linear algebra is a theory for solving linear equations.

3.2.2 Systems of equations

Now we consider a system of linear equations in two variables x_1, x_2.

$$a_{11}x_1 + a_{12}x_2 = b_1 \tag{1}$$
$$a_{21}x_1 + a_{22}x_2 = b_2\,. \tag{2}$$

We will see later that this corresponds to

$$(a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2 \tag{3}$$
$$(a_{11}a_{22} - a_{12}a_{21})x_2 = -a_{21}b_1 + a_{11}b_2 \tag{4}$$

Now we have two equations, one for x_1 and one for x_2. We know that the system has infinitely many solutions if and only if the following three equations hold

$$\begin{aligned} a_{11}a_{22} - a_{12}a_{21} &= 0 \\ a_{22}b_1 - a_{12}b_2 &= 0 \\ -a_{21}b_1 + a_{11}b_2 &= 0 \end{aligned}$$

The system has no solution if and only if

$$a_{11}a_{22} - a_{12}a_{21} = 0$$

and

$$(a_{22}b_1 - a_{12}b_2 \neq 0 \qquad \text{or} \qquad -a_{21}b_1 + a_{11}b_2 \neq 0)$$

hold.
The system has a unique solution if and only if

$$a_{11}a_{22} - a_{12}a_{21} \neq 0\,.$$

Of course we can continue that way for 3 equations in 3 variables and n equations in n variables.

Example. We want to solve

$$\begin{aligned} 3x_1 + x_2 &= 2 \\ 5x_1 + 2x_2 &= 3 \end{aligned}$$

We first eliminate x_1 in the second row:

$$\begin{aligned} 3x_1 + x_2 &= 2 \\ 3x_1 + \frac{6}{5}x_2 &= \frac{9}{5} \end{aligned}$$

$$\begin{aligned} 3x_1 + x_2 &= 2 \\ 0x_1 + \frac{1}{5}x_2 &= -\frac{1}{5} \end{aligned}$$

$$\begin{aligned} 3x_1 + x_2 &= 2 \\ x_2 &= -1 \end{aligned}$$

and then x_2 in the first row:

$$\begin{aligned} 3x_1 \qquad &= 3 \\ x_2 &= -1 \end{aligned}$$

and get

$$\begin{aligned} x_1 &= 1 \\ x_2 &= -1\,. \end{aligned}$$

3.3 Matrices

3.3.1 Vector spaces

What is the meaning of what we have done?
We consider a linear function

$$\begin{array}{rccl} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & f(x) := ax\,, \end{array}$$

$a \in \mathbb{R}$. Then the equation $ax = b$ has a solution if and only if b belongs to the image of f, i.e. if there is $x \in \mathbb{R}$ with $f(x) = b$.
For the system of two equations in two variables we consider a linear function

$$\begin{array}{rccl} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & x & \longmapsto & f(x) := Ax\,. \end{array}$$

with

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$$

a vector and

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

a 2×2-matrix. With this notation the system of equations can be written as

$$Ax = b\,,$$

where

$$b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2\,.$$

We write

$$Ax = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$$

and

$$Ax = b \quad \Longleftrightarrow \quad \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\,.$$

This is equivalent to

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

This can be written as

$$\sum_{j=1}^{2} a_{ij}x_j = b_i\,, \quad i = 1, 2\,.$$

The system

$$\begin{aligned} 3x_1 + x_2 &= 2 \\ 5x_1 + 2x_2 &= 3 \end{aligned}$$

can be written

$$Ax = b$$

with

$$A := \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}, \quad b := \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad x := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

The solution is

$$x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

A system of m equations in n variables

$$\sum_{j=1}^{n} a_{ij} x_j = b_i\,, \quad i = 1, \dots m\,.$$

can be written as

$$Ax = b\,,$$

with

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

a $m \times n$-matrix,

$$x := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

a vector in the vector space $\mathbb{R}^n$ and

$$b := \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$$

a vector in the vector space $\mathbb{R}^m$.

The product Ax equals

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$
$$= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}$$

The system of equations can be written

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} .$$

This is equivalent to

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{2m}x_n &= b_m . \end{aligned}$$

Here we have a linear function

$$\begin{aligned} f: \ \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x &\longmapsto f(x) := Ax . \end{aligned}$$

The equation $Ax = b$ has a solution if and only if $b \in \mathbb{R}^m$ belongs to the image of the function f, i.e. if there is a $x \in \mathbb{R}^n$ with $f(x) = b$.

Definition 3.3. A $n \times n$ matrix

$$A := (a_{ij})_{\substack{i=1,\ldots,n \\ j=1,\ldots,n}}$$

is called an *upper triangular* matrix if and only if

$$a_{ij} = 0 \text{ for } i > j$$

and *lower triangular* matrix if and only if

$$a_{ij} = 0 \text{ for } i < j\,.$$

The matrix

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}$$

is upper triangular and the matrix

$$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n1} & \dots & a_{n\,(n-1)} & a_{nn} \end{pmatrix}$$

is a lower triangular matrix.

3.3.2 Linear dependency

Definition 3.4. We consider a set of n vectors $\alpha_1, \dots \alpha_n \in \mathbb{R}^m$

$$\alpha_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}\,, \quad \alpha_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}\,, \quad \dots\,, \quad \alpha_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}\,.$$

This set is called *linearly dependent* if there exist $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ such that

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_n\alpha_n = 0\,.$$

and there is $i \in \{1, \dots n\}$ with $\lambda_i \neq 0$ (at least one of the λ_i is not 0). The set of vectors $\alpha_1, \dots \alpha_n \in \mathbb{R}^m$ is called *linearly independent* if and only if

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_n\alpha_n = 0$$

implies

$$\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0\,.$$

Example. The set

$$\alpha_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\,, \quad \alpha_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\,, \quad \alpha_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

is linearly independent. Indeed

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \lambda_3\alpha_3 = \lambda_1 \begin{pmatrix}1\\0\\1\end{pmatrix} + \lambda_2 \begin{pmatrix}0\\1\\0\end{pmatrix} + \lambda_3 \begin{pmatrix}1\\1\\-1\end{pmatrix}$$
$$= \begin{pmatrix}\lambda_1+\lambda_3\\ \lambda_2+\lambda_3\\ \lambda_1-\lambda_3\end{pmatrix} \overset{!}{=} \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

This implies $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

Definition 3.5. The set

$$< \alpha_1, \ldots, \alpha_n > := \left\{ \lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_n\alpha_n \,\middle|\, \lambda_1, \ldots, \lambda_n \in \mathbb{R} \right\}$$

is called the *span* of $\alpha_1, \ldots, \alpha_n$. It is the set of all linear combinations of $\alpha_1, \ldots, \alpha_n$. The dimension of $V := < \alpha_1, \ldots, \alpha_n >$ is the maximal number m such that there are sets of m linearly independent vectors in V. Such a set is called a *basis* of V.

Remark 3.6. If $\{\alpha_1, \ldots, \alpha_n\}$ is a basis of the n-dimensional vector space V, then for every vector $x \in V$ there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that

$$x = \sum_{i=1}^{n} \lambda_i\alpha_i = \lambda_1\alpha_1 + \ldots + \lambda_n\alpha_n .$$

The dimension of $\mathbb{R}^n$ is n and the unit-vectors

$$e_1 = \begin{pmatrix}1\\0\\ \vdots\\ \vdots\\0\end{pmatrix},\ e_2 = \begin{pmatrix}0\\1\\0\\ \vdots\\0\end{pmatrix}, \ldots,\ e_n = \begin{pmatrix}0\\0\\ \vdots\\0\\1\end{pmatrix} \tag{5}$$

form a basis of $\mathbb{R}^n$. It is called the *standard basis* of $\mathbb{R}^n$. The vector

$$\alpha = \begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}$$

is written in the standard basis of $\mathbb{R}^n$ and

$$\alpha = \begin{pmatrix}a_1\\a_2\\ \vdots\\ \vdots\\a_n\end{pmatrix} = a_1 \begin{pmatrix}1\\0\\ \vdots\\ \vdots\\0\end{pmatrix} + a_2 \begin{pmatrix}0\\1\\0\\ \vdots\\0\end{pmatrix} + \ldots + a_n \begin{pmatrix}0\\0\\ \vdots\\0\\1\end{pmatrix} = \sum_{i=1}^{n} a_i e_i .$$

Let A be the matrix whose columns are the vectors $\alpha_1, \dots, \alpha_n$, i.e.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \alpha_1, & \alpha_2, & \dots, & \alpha_n \end{pmatrix}$$

Then

$$\alpha_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$$

is the image of e_i under the mapping

$$\begin{array}{rccl} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & x & \longmapsto & f(x) := Ax \end{array}$$

and the image of $\mathbb{R}^n$ under f is

$$\begin{aligned} \operatorname{im}(f) &:= \left\{ y \in \mathbb{R}^m \,\middle|\, \exists x \in \mathbb{R}^n \,,\ f(x) = Ax = y \right\} \\ &= < \alpha_1, \dots, \alpha_n > \ . \end{aligned}$$

Therefore the dimension of the image is the dimension of $< \alpha_1, \dots, \alpha_n >$. Moreover $\dim(\operatorname{im}(f)) \leq n$.

3.3.3 Addition and product of matrices

Given two $m \times n$-matrices $A = (a_{ij})$ and $B = (b_{ij})$ we define the sum $A + B$ to be

$$A + B = (a_{ij} + b_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} .$$

Let $\alpha \in \mathbb{R}$ be a scalar then

$$\alpha A = (\alpha \, a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} .$$

Definition 3.7. The product of the $m \times l$-matrix

$$A = (a_{ik})_{\substack{i=1,\dots,m \\ k=1,\dots,l}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1l} \\ a_{21} & a_{22} & \cdots & a_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{ml} \end{pmatrix}$$

and the $l \times n$-matrix

$$B = (b_{kj})_{\substack{k=1,\dots,l \\ j=1,\dots,n}} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{l1} & b_{l2} & \cdots & b_{ln} \end{pmatrix}$$

is the $m \times n$-matrix $AB = C$

$$(c_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = C = AB = \Big(\sum_{k=1}^{l} a_{ik} b_{kj} \Big)_{\substack{i=1,\dots,m \\ j=1,\dots,n}} .$$

The component c_{ij} in AB is given by the scalar product of the i-th row of A with the j-th column of B.

Remark 3.8. In general

$$A\,B \neq B\,A\,.$$

The $n \times n$-identity is the matrix $\mathrm{Id}_n = (c_{ij})_{i,j=1,\dots,n}$ with

$$\mathrm{Id}_n\, A = A\, \mathrm{Id}_n = A\,.$$

We have

$$c_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j\,, \end{cases}$$

$$\mathrm{Id}_n := \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} .$$

Given a $n \times n$-matrix A. If there is a $n \times n$-matrix B with $BA = \mathrm{Id}_n$, then

$$AB = BA = \mathrm{Id}_n$$

and $B = A^{-1}$ is called the *inverse* of A.
A method to compute the inverse of a matrix is the following. It is much easier to show on an example than to explain. Given a matrix

$$\begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$$

that we want to invert. We write

$$\begin{array}{c|c}
\begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\begin{pmatrix} 3 & 1 \\ 3 & \frac{6}{5} \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & \frac{3}{5} \end{pmatrix} \\
\begin{pmatrix} 3 & 1 \\ 0 & \frac{1}{5} \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ -1 & \frac{3}{5} \end{pmatrix} \\
\begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ -5 & 3 \end{pmatrix} \\
\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 6 & -3 \\ -5 & 3 \end{pmatrix} \\
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}
\end{array}$$

Now the inverse of

$$\begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$$

is

$$\begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix} .$$

3.3.4 The system of equations

We consider the equation $Ax = b$ with

$$A := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} , \qquad x := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} , \qquad b := \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} .$$

Then we multiply

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

from the left with $\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$ and get

$$\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

$$\begin{pmatrix} a_{11}a_{22} - a_{12}a_{21} & 0 \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{22}b_1 - a_{12}b_2 \\ -a_{21}b_1 + a_{11}b_2 \end{pmatrix} .$$

This is equivalent to

$$\begin{aligned}(a_{11}a_{22} - a_{12}a_{21})x_1 &= a_{22}b_1 - a_{12}b_2 \\ (a_{11}a_{22} - a_{12}a_{21})x_2 &= -a_{21}b_1 + a_{11}b_2\,.\end{aligned}$$

Herewith we get the results of section 3.2.2.

3.3.5 The transpose of a matrix

Definition 3.9. We define the *transpose* A^{t} of the matrix

$$A := (a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$$

to be the matrix

$$A^{\mathrm{t}} := (a_{ji})_{\substack{j=1,\dots,n\\ i=1,\dots,m}}\,.$$

It follows immediately that

$$(A^{\mathrm{t}})^{\mathrm{t}} = A\,.$$

The transpose of A is obtained in exchanging the rows with the columns. The i-th row of A is the i-th colum of A^{t} and the j-th column of A is the j-th row of A^{t}.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{\mathrm{t}} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

Given the column vector

$$a = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_i \\ \vdots \\ \alpha_n \end{pmatrix},$$

the transpose a^{t} of a is the row vector

$$a^{\mathrm{t}} = (\alpha_1\,,\ \dots\,,\ \alpha_i\,,\ \dots\,,\ \alpha_n)\,.$$

We define the scalar product of a and b,

$$b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_i \\ \vdots \\ \beta_n \end{pmatrix}$$

to be

$$\langle a, b\rangle := a^{\mathrm{t}}\, b = (\alpha_1\,,\;\ldots\;,\;\alpha_i\,,\;\ldots\;,\;\alpha_n)\begin{pmatrix}\beta_1\\ \vdots\\ \beta_i\\ \vdots\\ \beta_n\end{pmatrix}$$
$$= \alpha_1\beta_1 + \cdots + \alpha_i\beta_i + \cdots + \alpha_n\beta_n$$
$$= \sum_{i=1}^{n} \alpha_i\beta_i\,.$$

3.3.6 Determinant

Let A be a $n \times n$-matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix} = \begin{pmatrix}\alpha_1, & \alpha_2, & \ldots, & \alpha_n\end{pmatrix}$$

The determinant $\det(A)$ of a $n \times n$-matrix A is a function that maps the matrix, the set of its n column-vectors to an element of the underlying field ($\mathbb{R}$). This function satisfies the following properties for $1 \leq i, j \leq n$

i) For $a \in \mathbb{R}$

$$\det(\alpha_1, \ldots, a\,\alpha_i, \ldots, \alpha_n) = a \cdot \det(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n)$$

ii) With $\alpha_i = \alpha'_i + \alpha''_i$

$$\det(\alpha_1, \ldots, \alpha'_i + \alpha''_i, \ldots, \alpha_n)$$
$$= \det(\alpha_1, \ldots, \alpha'_i, \ldots, \alpha_n) + \det(\alpha_1, \ldots, \alpha''_i, \ldots, \alpha_n)\,.$$

iii) Let $\alpha_i = \alpha_j$ for $i \neq j$. Then

$$\det(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_j, \ldots, \alpha_n) = 0\,.$$

iv) Let $e_1 \ldots, e_n$ be the unit-vectors (5), then

$$\det(e_1, \ldots, e_n) = 1\,.$$

The determinant has the following important properties.

Theorem 3.10. *The determinant of a $n \times n$-matrix*

$$(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_n)$$

is 0 *if and only if the vectors $\alpha_1, \ldots, \alpha_n$ are linearly dependent.*
For the transpose of A we have

$$\det A = \det A^{\mathrm{t}}\,.$$

For the product of the $n \times n$-matrices A and B

$$\det(AB) = \det A \cdot \det B\,.$$

Let $\alpha \in \mathbb{R}$ be a scalar, then

$$\det(\alpha\, A) = \alpha^n \det A\,.$$

Example.

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc\,.$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}\,.$$

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{array}{l} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13} \\ -a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{23}a_{32}a_{11}\,. \end{array}$$

For $n \times n$-matrices there is an analogous formula for developing by columns or rows. For a $n \times n$-matrix $A = (a_{ij})$ we define A_{ij} to be the $(n-1) \times (n-1)$ that we get from A if we take off the ith row and the jth column.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}$$

$$A_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{(i-1)1} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)n} \\ a_{(i+1)1} & \cdots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \cdots & a_{(i+1)n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}$$

Then developing by the jth column

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

or by the ith row

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij} \,.$$

For a 3×3-matrix we get if we develop by the first column

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= (-1)^{1+1} a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \\ &\quad + (-1)^{2+1} a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} \\ &\quad + (-1)^{3+1} a_{31} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} \,. \end{aligned}$$

In analogy one can develop by any other column or a row. With the second row we get

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= (-1)^{2+1} a_{21} \det \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} \\ &\quad + (-1)^{2+2} a_{22} \det \begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix} \\ &\quad + (-1)^{2+3} a_{23} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix} \,. \end{aligned}$$

Example. We compute the determinant

$$\det \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

Since the second column contains two zeros, we choose

$$\det \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix} = (-1)^{2+2} \det \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -1 - 1 = -2 \,.$$

3.4 Eigenvalues and eigenvectors

3.4.1 Change of basis

We defined the standard-basis of $\mathbb{R}^n$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \; e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

We know that a maximal set of linearly independent vector $\tau_1, \ldots, \tau_n$ in a vector space is called a basis.
We know that in the jth column of A is the image of e_j under

$$\begin{array}{rccl} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & x & \longmapsto & f(x) := Ax\,. \end{array}$$

This implies that the matrix A corresponding to f depends on the choice of the basis of $\mathbb{R}^n$. Let $\tau_1, \ldots, \tau_n$ be another basis of $\mathbb{R}^n$. It is possible to express the mapping f in the basis $\tau_1, \ldots, \tau_n$ by a matrix A'. The jth column α'_j of A' is the image of τ_j written as a linear combination of the $\tau_1, \ldots, \tau_n$:

$$\alpha'_j = \sum_{i=1}^{n} a'_{ij} \tau_i\,.$$

Let τ_j be written in the basis $e_1, \ldots, e_n$, i.e.

$$\tau_j = \sum_{i=1}^{n} t_{ij} e_i = \begin{pmatrix} t_{1j} \\ \vdots \\ t_{nj} \end{pmatrix}.$$

Let T be the matrix with column-vectors $\tau_1, \ldots, \tau_n$, i.e.

$$T = (\tau_1\,, \ldots\,, \tau_n) = \begin{pmatrix} t_{11} & \cdots & t_{1j} & \cdots & t_{1n} \\ \vdots & & \vdots & & \vdots \\ t_{i1} & \cdots & t_{ij} & \cdots & t_{in} \\ \vdots & & \vdots & & \vdots \\ t_{n1} & \cdots & t_{nj} & \cdots & t_{nn} \end{pmatrix}.$$

Since $\tau_1, \ldots, \tau_n$ are a basis and herewith linearly independent, we know that T is invertible and $\det T \neq 0$. Then the matrix A' that describes the mapping f in the basis $\tau_1, \ldots, \tau_n$ is given by

$$A' = T^{-1} A\, T$$

and herewith

$$T\,A'\,T^{-1} = A\,.$$

Here we use that in general $AB \neq BA$.

Theorem 3.11. *The determinant is invariant under the change of basis.*

Proof. Since $\det(AB) = \det A \, \det B$ we have

$$\det A' = \det(T^{-1}\,A\,T) = \det T^{-1}\, \det A\, \det T = \det A$$

where the last equation follows from

$$1 = \det \mathrm{Id}_n = \det T^{-1}\, \det T$$

which is equivalent to

$$\det T^{-1} = \frac{1}{\det T}\,.$$

□

Example. In the standard basis e_1, e_2, e_3 the mapping

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

is given by

$$x := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto Ax := \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

We choose a new basis

$$\tau_1 := e_1 - e_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \tau_2 := e_2 - e_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \tau_3 := e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Then the transformation from the standard basis to the basis $\{\tau_1, \tau_2, \tau_3\}$ is given by the matrix

$$T := \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

We have $\det T = 1$ and in the basis $\{\tau_1, \tau_2, \tau_3\}$ the mapping f is given by

$$x := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto T^{-1} A T\, x\,.$$

Here

$$T^{-1} := \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

and

$$\begin{aligned} T^{-1}AT &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 0 & -2 & 4 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} . \end{aligned}$$

3.4.2 Scalings

We know that a linear mapping maps the vector space $\mathbb{R}^n$ to a subspace of $\mathbb{R}^m$. In this section we study some examples of linear mappings.
The first example is the scaling. The scaling with a factor λ along the direction $< v >$ maps v to λv. Given a mapping

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

If there exist n linearly independent vectors $v_1, \ldots, v_n$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that

$$f(v_i) = \lambda_i v_i \,, \qquad i = 1, \ldots n,$$

then f acts on $< v_i >$ as a scaling by the factor λ_i.
In the basis $v_1, \ldots, v_n$ the matrix representation of f is

$$A' = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} .$$

Let

$$T = (\, v_1 \,, \, \ldots \,, \, v_n \,)$$

be the matrix with columns $v_1, \ldots, v_n$. Then the gives the change of basis from the standard basis to $v_1, \ldots, v_n$. In the standard basis the mapping f is given by the matrix

$$A = T\, A'\, T^{-1} \,.$$

In the next section we see how to find for a given matrix A the vectors $v_1, \ldots, v_n$ and the factors $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

3.4.3 Eigenvalues and eigenvectors

Given a matrix A we consider the equation

$$Av = \lambda v\,,$$

i.e. we search for factors $\lambda \in \mathbb{R}$ and vectors $v \in \mathbb{R}^n$ such that the equation $Av = \lambda v$ holds.
We have

$$A\,v = \lambda\,v \quad \Longleftrightarrow \quad (A - \lambda\,\mathrm{Id}_n)\,v = 0\,.$$

Of course this is always true for $v = 0$. There exists $v \neq 0$ with $(A - \lambda\,\mathrm{Id}_n)\,v = 0$ if and only if $\det(A - \lambda \mathrm{Id}_n) = 0$.

Definition 3.12. The *eigenvalues* of a matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ are the scalars λ that satisfy the equation

$$\det(A - \lambda\,\mathrm{Id}_n) = 0\,.$$

A vector that satisfies the equation

$$(A - \lambda\,\mathrm{Id}_n)\,v = 0$$

for a given matrix A and an eigenvalue λ of A, is called *eigenvector* of A to the eigenvalue λ.

The eigenvalues of the $n \times n$-matrix A are the solutions of the equation

$$p_A(\lambda) = \det(A - \lambda\,\mathrm{Id}_n) = 0\,.$$

This equation is a polynomial in λ of degree n.

Definition 3.13. Given a matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$, the polynomial

$$p_A(\lambda) = \det(A - \lambda\,\mathrm{Id}_n) = 0\,.$$

is called the *characteristic polynomial* of A.

Definition 3.14. The *algebraic multiplicity* of the eigenvalue λ_i of A is the k that is minimal with

$$(\lambda - \lambda_i)^k \text{ is a factor of } p_A(\lambda) = \det(A - \lambda\,\mathrm{Id}_n)\,.$$

The *geometric multiplicity* is the dimension of the subspace spanned by the eigenvectors of A to the eigenvalue λ.

Remark 3.15. The geometric multiplicity of an eigenvalue is less or equal to the algebraic multiplicity of the eigenvalue.

4 Multivariable Calculus: second part

4.0.4 Local extrema

Definition 4.1. Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ a function. A point $x \in U$ is called local maximum, resp. local minimum, of f if an environment $V \subset U$ of x exists with

$$f(x) \geq f(y), \quad \big(\text{resp. } f(x) \leq f(y)\big) \quad \text{for all } y \in V\,.$$

Definition 4.2. Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$, $(x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n)$ a function whose first and second partial derivatives exist and are continuous. The Hessian matrix of f in $x \in U$ is the $n \times n$-matrix

$$(\operatorname{Hess} f)(x) := \left(\frac{\partial^2}{\partial x_i \partial x_j} f(x) \right)_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} .$$

This matrix is symmetric since for $1 \leq i \leq n$, $1 \leq j \leq n$

$$\frac{\partial^2}{\partial x_i \partial x_j} f(x) = \frac{\partial^2}{\partial x_j \partial x_i} f(x)\,.$$

Theorem 4.3. *Let $U \subset \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ a partial differentiable function. If f has a local Extremum in the point x (i.e. a local maximum or a local minimum), then*

$$\nabla f(x) = 0\,.$$

Theorem 4.4. *Let $U \subset \mathbb{R}^n$ be open, $f : U \to \mathbb{R}$ a function whose first and second partial derivatives exist and are continuous. Let $x \in U$ with*

$$\nabla f(x) = 0\,.$$

Then

- *if $(\operatorname{Hess} f)(x)$ is positive definite, then f has a local minimum in x.*
- *if $(\operatorname{Hess} f)(x)$ is negative definite, then f has a local maximum in x.*
- *if $(\operatorname{Hess} f)(x)$ is indefinite, then f doesn't have a local extremum in x.*

In the other cases there may not be a local extremum.

Definition 4.5. A symmetric matrix A is

- positive definite if all its eigenvalues are positive,

- negative definite if all its eigenvalues are negative,
- indefinite if there is at least one positive and one negative eigenvalue.

Example. i) The function $f(x, y) := c + x^2 + y^2$ has in $(0, 0)$ a local minimum since $\nabla f(0, 0) = (0, 0)$ and the Hessian

$$(\mathrm{Hess}\, f)(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

is positive definite.

ii) The function $g(x, y) := c - x^2 - y^2$ has in $(0, 0)$ a local maximum since $\nabla g(0, 0) = (0, 0)$ and the Hessian

$$(\mathrm{Hess}\, g)(0,0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

is negative definite.

iii) The function $h(x, y) := c + x^2 - y^2$ satisfies $\nabla h(0, 0) = (0, 0)$ and the Hessian

$$(\mathrm{Hess}\, h)(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

is indefinite. It has a saddle point in $(0, 0)$.

4.1 Surfaces

4.1.1 Parameterized surfaces

Definition 4.6. Let D be a subset of $\mathbb{R}^2$ (i.e. a rectangle). Then the function

$$\begin{array}{rccl} \vec{r}: & D & \longrightarrow & \mathbb{R}^3 \\ & (u,v) & \longmapsto & \vec{r}(u,v) := \big(x(u,v), y(u,v), z(u,v)\big) \end{array}$$

is a parameterization of a surface S.

We consider a point $(u_0, v_0) \in D$ and its image $p := \vec{r}(u_0, v_0)$ on S. We keep v_0 fixed and consider the partial function

$$u \longmapsto \vec{r}(u, v_0)\,.$$

This is a curve on S through p, the u-line. In p the u-line has the tangent vector

$$\vec{r}_u(u_0, v_0) =: \vec{r}_u\,.$$

In analogy we have the v-line

$$v \longmapsto \vec{r}(u_0, v)$$

on S with tangent vector

$$\vec{r}_v(u_0, v_0) =: \vec{r}_v\,.$$

The parameter representation

$$\vec{r} : D \longrightarrow \mathbb{R}^3$$

is regular in (u_0, v_0) if the Jacobian matrix

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{pmatrix}$$

has rank 2 (i.e. the 2 column vectors are linearly independent) or equivalently if the cross product (also called vector product)

$$\vec{r}_u \times \vec{r}_v \neq 0\,.$$

The vector

$$\vec{n} := \frac{\vec{r}_u \times \vec{r}_v}{|\vec{r}_u \times \vec{r}_v|}\,.$$

is perpendicular to the tangent space of S in p. It is the normal unit vector to S in p.

4.1.2 Coordinates

Change of the coordinate system can be considered as mappings

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \qquad f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3\,.$$

For example the change of from cartesian to sperical coordinates is given by

$$f : \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ (r, \theta, \varphi) & \longmapsto & (x, y, z) \end{array}$$

with

$$\begin{aligned} x &= r\,\sin\theta\,\cos\varphi \\ y &= r\,\sin\theta\,\sin\varphi \\ z &= r\,\cos\theta \end{aligned}$$

We consider the function

$$\begin{array}{rccc} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & (x_1,\ldots,x_n) & \longmapsto & f(x_1,\ldots,x_n) \end{array}$$

with

$$f(x_1,\ldots,x_n) := \big(f_1(x_1,\ldots,x_n),\ldots,f_m(x_1,\ldots,x_n)\big)\,.$$

Here the functions f_i, $i = 1,\ldots,m$ are scalar fields. We suppose that the partial derivatives of these functions exist.

Definition 4.7. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be given as above. We define the Jacobian matrix of f to be $n \times m$-matrix

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

This is a linear mapping between tangent spaces and

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} \partial x_1 \\ \vdots \\ \partial x_n \end{pmatrix} = \begin{pmatrix} \partial f_1 \\ \vdots \\ \partial f_m \end{pmatrix}.$$

In the example of spherical coordinates the Jacobian is

$$\begin{pmatrix} \sin\theta\,\cos\varphi & r\,\cos\theta\,\cos\varphi & -r\,\sin\theta\,\sin\varphi \\ \sin\theta\,\sin\varphi & r\,\cos\theta\,\sin\varphi & r\,\sin\theta\,\cos\varphi \\ \cos\theta & -r\,\sin\theta & 0 \end{pmatrix}.$$

4.1.3 The area of a surface

In order to determine the area $\omega(S)$ of the surface S, we consider the area of the image of a small rectangle $D_k \subseteq D$, $1 \le k \le N$:

$$D_k := [u_k, u_k + \Delta u] \times [v_k, v_k + \Delta v]\,.$$

The area of this parallelogram is

$$\omega(S_k) = \big|\vec{r}_u(u_k, v_k) \times \vec{r}_v(u_k, v_k)\big|\,\mu(D_k)\,,$$

where $\mu(D_k)$ is the volume of D_k. The area of S is

$$\omega(S) := \int_D \big|\vec{r}_u(u, v) \times \vec{r}_v(u, v)\big|\,d\mu(u, v)\,.$$

Where

$$d\omega := \big|\vec{r}_u(u,v) \times \vec{r}_v(u,v)\big|\, d\mu(u,v)$$

is the *(scalar) surface element*.
In cartesian coordinates

$$d\mu(u,v) = du\, dv\,.$$

The transformation to polar coordinates is given by

$$x = r\,\cos\varphi, \qquad y = r\,\sin\varphi\,.$$

The corresponding Jacobian is

$$J = \begin{pmatrix} \cos\varphi & -r\,\sin\varphi \\ \sin\varphi & r\,\cos\varphi \end{pmatrix}\,.$$

Then

$$d\mu(r,\varphi) = \det(J)\, dr\, d\varphi = r\, dr\, d\varphi\,.$$

Example. The area of the 2-sphere of radius R is given by

$$\begin{aligned}\omega(S^2_R) &= \int_{[0,2\pi]\times[-\pi/2,\pi/2]} |\vec{r}_\varphi \times \vec{r}_\theta| d\mu(\varphi,\theta) \\ &= \int_0^{2\pi} \Big(\int_{-\pi/2}^{\pi/2} R^2 \cos\theta d\theta\Big) d\varphi = \int_0^{2\pi} 2R^2 d\varphi \\ &= 4\pi R^2\,.\end{aligned}$$

Example. The hyperbolic paraboloid is given by the equation

$$z = x^2 - y^2\,.$$

Let S be the part of the surface that is in the cylinder $x^2 + y^2 \le R^2$. To compute the area $\omega(S)$ we first consider the surface element of a graph

$$S: \quad (x,y) \mapsto \vec{r}(x,y) = \big(x, y, f(x,y)\big)\,.$$

We have

$$\vec{r}_x = (1,0,f_x), \quad \vec{r}_y = (0,1,f_y)$$

and herewith

$$\vec{r}_x \times \vec{r}_y = (-f_x, -f_y, 1)\,.$$

This yields the surface element

$$d\omega = \sqrt{1 + f_x^2 + f_y^2}\, d\mu(x,y)\,.$$

For the example $f(x, y) := x^2 - y^2$ we have

$$\omega(S) = \int_B d\omega = \int_B \sqrt{1 + 4x^2 + 4y^2}\, d\mu(x, y)\,,$$

where B is the disc of radius R in the (x, y)-plane. In polar coordinates we have

$$\begin{aligned}\omega(S) &= \int_0^{2\pi} \Big(\int_0^R \sqrt{1 + 4r^2}\, r\, dr \Big) d\varphi \\ &= 2\pi \cdot \frac{1}{12} (1 + 4r^2)^{3/2} \Big|_0^R \\ &= \frac{\pi}{6} \big((1 + 4R^2)^{3/2} - 1 \big)\,.\end{aligned}$$

5 Differential equations

5.1 Definition

Definition 5.1. A *differential equation* is an equation

$$F\big(x, y(x), y'(x), \ldots, y^{(n)}(x)\big) = 0$$

relating the independent variable x, the unknown function

$$y : \mathbb{R} \longrightarrow \mathbb{R}\,.$$

and its derivatives.
The ith derivative of y is denoted by $y^{(i)}$.
The *order* of the differential equation is the order of the highest derivative that appears in the relation F. (Here it is $y^{(n)}$ and therefore the order is n.)
The differential equation is called *ordinary* if y is a function of a single variable. If $y : \mathbb{R}^n \to \mathbb{R}$ is a function in more than one variable, then the differential equation is called a *partial* differential equation.

5.2 First order differential equations

Definition 5.2. A differential equation of order 1 has the following general form:

$$y'(x) = f\big(x, y(x)\big)\,,$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ with $\operatorname{dom}(f) = D \subseteq \mathbb{R}^2$.

The solution y of a differential equation of order one is a curve that has a given slope in every point.

Theorem 5.3. *The following statements are true under the assumption of some technical conditions for the function f.*

i) The solutions of a differential equation $y'(x) = f(x, y(x))$ form a one-parameter family of curves $y_c : x \to y_c(x)$.

ii) The initial value problem (IVP)

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0 \end{cases}$$

has a unique solution $x \mapsto y(x)$.

5.2.1 Separation of variables

Some differential equations are of the form

$$y'(x) = f(y) \cdot g(x)\,.$$

These have the following properties:

- The extrema of $y(x)$ are in the zeros of $g(x)$.
- The zeros of $f(y)$ are solutions since for y_0 with $f(y_0) = 0$ holds:

$$\text{Let } y(x) = y_0, \text{ then } y'(x) = 0 = f(y_0) \cdot g(x).$$

W.l.o.g. we assume $f(y) \neq 0$ and define

$$h(y) = \frac{1}{f(y)}\,.$$

We separate

$$\begin{aligned} \frac{dy}{dx} &= f(y) \cdot g(x) \\ h(y)\,dy &= g(x)\,dx\,. \end{aligned}$$

Integration yields

$$\int h(y)\,dy = \int g(x)\,dx\,.$$

Let $H(y)$ be an antiderivative of $h(y)$ and $G(x)$ an antiderivative of $g(x)$. Let $y = y(x)$. Then for all x on an interval (a, b) the equation

$$H(y(x)) = G(x)$$

has to be satisfied. We compute the derivative with respect to x:

$$\frac{dH}{dy} \cdot \frac{dy}{dx} = \frac{dG}{dx}\,,$$

i.e.

$$h(y) \cdot y' = g(x)\,.$$

This is equivalent to our differential equation.
The solution of the initial value problem

$$\begin{cases} y' &= g(x)f(y) \\ y(x_0) &= y_0 \end{cases}$$

is given by

$$\int_{y_0}^{y(x)} \frac{ds}{f(s)} = \int_{x_0}^{x} g(t)\,dt\,.$$

Example. The differential equation

$$y' = \frac{-x}{y}$$

is separable:

$$\int y\,dy = -\int x\,dx\,.$$

The solution is a curve in $\mathbb{R}^2$ with slope

$$f(x,y) = \frac{-x}{y}$$

in any point $(x,y) \in \mathbb{R}^2$ in which f is defined. The solution satisfies

$$y^2 = -x^2 + c^2$$

i.e.

$$y = \sqrt{c^2 - x^2}$$

The constant c is defined by the initial value (x_0, y_0).

Type $\mathbf{y' = f(ax + by + c)}$ We consider the differential equation

$$y' = f(ax + by + c)$$

and introduce the unknown function

$$u(x) = ax + by(x) + c\,.$$

Then

$$u' = a + by'\,,$$

i.e.

$$y' = \frac{u' - a}{b}\,,$$

We get a differential equation for u

$$\frac{u' - a}{b} = f(u) \Longleftrightarrow u' = a + b \cdot f(u)\,.$$

This equation is separable:

$$\int \frac{du}{a + b \cdot f(u)} = x\,.$$

If it is possible to compute this integral $(= I(u))$, the equation

$$I(ax + by + c) = x$$

has to be solved with respect to y (i.e. $y = \ldots$).

Example. We solve the initial value problem

$$\begin{cases} y' + y &= 1 + x,\ x > 0 \\ y(0) &= 2\,. \end{cases}$$

The differential equation is

$$y' = x - y + 1\,.$$

We define

$$u = x - y + 1$$

and have

$$u' = 1 - y'$$

i.e.

$$y' = 1 - u'\,.$$

The differential equation is

$$1 - u' = f(u) \Longleftrightarrow u' = 1 - f(u)$$

with $f(u) = u$. The equation

$$u' = 1 - u$$

is separable:

$$\begin{aligned} \int \frac{1}{1-u}\,du &= \int dx \\ -\ln(1-u) &= x + c_1 \\ \ln(-x+y) &= -x - c_1 \\ -x + y &= e^{-x-c_1} \\ y &= x + Ce^{-x}\,. \end{aligned}$$

The constant C is defined by the initial value $y(0) = 2$. The solution of the problem is

$$y = x + 2e^{-x}\,.$$

Type $\mathbf{y}' = \mathbf{f}\left(\frac{\mathbf{y}}{\mathbf{x}}\right)$ The differential equation

$$y' = f\left(\frac{y}{x}\right)$$

is also solved by substitution and separation. We choose the unknown function

$$u = \frac{y(x)}{x}\,.$$

Then

$$y' = (u \cdot x)' = u' \cdot x + u\,,$$

and the differential equation for u is

$$u' \cdot x + u = f(u) \Longleftrightarrow u' = \frac{f(u) - u}{x}\,.$$

Herewith

$$\underbrace{\int \frac{du}{f(u) - u}}_{I(u)} = \int \frac{dx}{x} = \ln|x| + \text{const} = \ln(c \cdot |x|)\,.$$

Sometimes it is possible to solve

$$I\left(\frac{y}{x}\right) = \ln(c \cdot |x|)$$

with respect to y.

5.3 Linear differential equations

Definition 5.4. A *linear differential equation* is a differential equation

$$F\big(y(x), y'(x), \ldots, y^{(n)}(x)\big) = q(x)$$

in which the function F is linear in y and the derivatives $y^{(i)}$ of y.
The differential equation is called *inhomogeneous* if $q(x) \neq 0$ and it is called *homogeneous* if $q(x) = 0$, i.e. if

$$F\big(y(x), y'(x), \ldots, y^{(n)}(x)\big) = 0$$

5.3.1 Variation of constants

We consider the following linear inhomogeneous differential equation of first order:

$$y'(x) + p(x)y(x) = q(x)\,.$$

First we consider the corresponding homogeneous differential equation:

$$y'(x) + p(x)y(x) = 0\,.$$

This equation is separable. We get

$$\frac{dy}{dx} = -p(x)y(x)$$
$$\int \frac{1}{y}\,dy = \int -p(x)\,dx$$

and herewith

$$y(x) = C \cdot e^{-P(x)}\,,$$

where

$$P(x) = \int_{x_0}^{x} p(x)dx$$

is an antiderivative of $p(x)$. We call $y(x) = C \cdot e^{-P(x)}$ the general solution of the homogeneous equation.

Remark 5.5. If $y_1(x)$ and $y_2(x)$ are two solutions of a homogeneous differential equation, then so is $\lambda\, y_1(x) + \mu\, y_2(x)$, $\lambda, \mu \in \mathbb{R}$.

To solve the inhomogeneous problem we make the following ansatz:

$$y(x) = C(x) \cdot e^{-P(x)}\,,$$

where the function $C(x)$ has to be defined. This is the variation of constants. With the differential equation

$$y'(x) + p(x)y(x) = q(x)$$

we get

$$C'(x) = q(x)\, e^{P(x)}$$

and herewith

$$C(x) = \int q(x)\, e^{P(x)}\, dx\,.$$

The particular solution of the inhomogeneous equation is

$$y(x) = e^{-P(x)} \int_{x_0}^{x} q(s)e^{P(s)}\, ds\,.$$

The following principle is true:

the general solution of an inhomogeneous differential equation
=
the general solution of the homogeneous differential equation
+
the particular solution of the inhomogeneous differential equation.

This yields in our case

$$y(x) = Ce^{-P(x)} + e^{-P(x)} \int_{x_0}^{x} q(s) e^{P(s)}\, ds\,.$$

The constant C is determined by the initial value $y(x_0) = y_0$.

Example. Find the maximal solution of the initial value problem

$$y' - y \cdot \tan x = \cos^2 x \quad \text{with} \quad y(0) = 1\,.$$

The solution of the homogeneous differential equation

$$y' - y \cdot \tan x = 0$$

is

$$y_h(x) = C \cdot e^{-P(x)}\,,$$

where $P(x)$ is an antiderivative of $p(x) = -\tan x$. We know that

$$-P(x) = \int_0^x \tan s\, ds = -\ln|\cos x| \text{ for } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \quad (x_0 = 0)$$

hence

$$y_h(x) = C \cdot e^{-\ln|\cos x|} = C \cdot \frac{1}{|\cos x|} \text{ for } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

The particular solution of the inhomogeneous equation is

$$\begin{aligned} y_p(x) &= e^{-P(x)} \int_{x_0}^{x} q(s) e^{P(s)}\, ds \\ &= \frac{1}{|\cos x|} \int_{x_0}^{x} \cos^2(s)\, |\cos s|\, ds\,, \end{aligned}$$

$x_0 = 0$. The general solution is

$$y(x) = C \cdot \frac{1}{|\cos x|} + \frac{1}{|\cos x|} \int_0^x \cos^2(s)\, |\cos s|\, ds$$

with $C = 1$ since $y_h(0) = 1$. Moreover $\cos x > 0$ for $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and

$$y(x) = \frac{1}{\cos x} + \frac{1}{\cos x} \int_0^x \cos^3(s)\, ds\,.$$

We compute

$$\begin{aligned}\int_0^x \cos^3(s)\, ds &= \int_0^x \cos(s)(1 - \sin^2 s)\, ds \\ &= \int_0^x \cos(s)\, ds - \int_0^x \cos(s) \sin^2 s\, ds \\ &= \sin x - \frac{1}{3} \sin^3 x\end{aligned}$$

and get

$$y(x) = \frac{1 + \sin x - \frac{1}{3}\sin^3 x}{\cos x} \quad \text{with } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

5.3.2 Linear differential equations of order n with constant coefficients

A linear differential equation of order n is

$$x^{(n)}(t) + a_{n-1}(t)x^{(n-1)}(t) + \ldots + a_1(t)\dot{x}(t) + a_0(t)x(t) = r(t)\,.$$

Remark 5.6.

- Linear means that there are no terms of the form $x^{(k)}(t) \cdot x^{(l)}(t)$.
- If $r(t) = 0$, then the differential equation is homogeneous. If $r(t) \neq 0$, then the differential equation is inhomogeneous.
- If $a_{n-1}(t), \ldots, a_0(t)$ are constant (don't depend on t), then the differential equation has constant coefficients.

We consider the homogeneous differential equation of order n with constant coefficients.

$$x^{(n)} + a_{n-1}x^{(n-1)} + \ldots + a_1\dot{x} + a_0 x = 0$$

The equation $\dot{x} = \lambda x$ has the general solution $x(t) = Ce^{\lambda t}$. We make the ansatz

$$x(t) = e^{\lambda t}\,, \quad \lambda \in \mathbb{R}$$

and get

$$\lambda^n e^{\lambda t} + a_{n-1}\lambda^{n-1}e^{\lambda t} + \ldots a_1\lambda e^{\lambda t} + a_0 e^{\lambda t} = 0$$

and

$$e^{\lambda t}(\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0) = 0\,.$$

Since $e^{\lambda t} \neq 0$, we have the characteristic polynomial

$$Q(\lambda) := \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0 = 0\,.$$

There are different cases:

i) $Q(\lambda)$ has n different real zeros.

ii) $Q(\lambda)$ has zeros with multiplicity > 1.

iii) $Q(\lambda)$ has complex zeros.

In the case i) the differential equation has n linear independent solutions:

$$e^{\lambda_1 t}, \ldots, e^{\lambda_n t}$$

and the general solution is

$$x(t) = \sum_{i=1}^{n} c_i e^{\lambda_i t}$$

with n constants $c_1, \ldots, c_n$.
Case ii): If λ_k is a zero with multiplicity p, then

$$e^{\lambda_k t}, te^{\lambda_k t}, t^2 e^{\lambda_k t}, \ldots, t^{p-1} e^{\lambda_k t}$$

are the p solutions corresponding to λ_k. Their contribution to the general equation is

$$C_0 e^{\lambda_k t} + C_1 t e^{\lambda_k t} + C_2 t^2 e^{\lambda_k t} + \cdots + C_{p-1} t^{p-1} e^{\lambda_k t}$$

Case iii): The simple complex zeros

$$\lambda = \alpha \pm i\beta$$

correspond to the real solutions

$$e^{\alpha t}\cos\beta t \quad \text{and} \quad e^{\alpha t}\sin\beta t\,.$$

These are linearly independent. If the mulitiplicity of the complex zeros is > 1, then we have the solutions

$$\begin{array}{llll} e^{\alpha t}\cos\beta t\,, & te^{\alpha t}\cos\beta t\,, & \ldots, & t^{p-1}e^{\alpha t}\cos\beta t\,, \\ e^{\alpha t}\sin\beta t\,, & te^{\alpha t}\sin\beta t\,, & \ldots, & t^{p-1}e^{\alpha t}\sin\beta t\,. \end{array}$$

Depending on the inhomogeneous term $r(t)$ there are different ansatz to make for the particular solution of the inhomogeneous equation. Here we give some of the most important examples

a) $r(t) = e^{at}$, $a \in \mathbb{R}$ or $a \in \mathbb{C}$. Then the ansatz is

$$x = c \cdot e^{at}$$

where c is a constant that has to be defined.

b) $r(t) = e^{\alpha t} \begin{cases} \sin \omega t \\ \cos \omega t \end{cases}$. This is similar to the previous case since

$$e^{\alpha t} \cdot \sin \omega t = \operatorname{Im} e^{(\alpha + i\omega)t}, \quad e^{\alpha t} \cdot \cos \omega t = \operatorname{Re} e^{(\alpha + i\omega)t}.$$

c) $r(t)$ is a polynomial in t of degree m. If $a_0 = \cdots = a_q = 0$, $a_{q+1} \neq 0$, then the ansatz is

$$x(t) = b_{m+1} t^{m+q+1} + \cdots + b_1 t^{1+q} + b_0 t^q.$$

Example. Determine all the solutions of the differential equation

$$y'' - y' - 6y = e^{-x}.$$

that are bounded on the interval $[0, \infty[$ and satisfy $y(0) = 0$.
We first consider the homogeneous equation

$$y'' - y' - 6y = 0.$$

The zeros of the characteristic polynomial

$$\lambda^2 - \lambda - 6 = 0$$

are

$$\lambda_{1,2} = \frac{1 \pm \sqrt{1+24}}{2} = \begin{cases} 3 \\ -2 \end{cases}.$$

The general solution of the homogeneous equation is

$$y_h(x) = c_1 e^{3x} + c_2 e^{-2x}.$$

The ansatz for the special solution of the inhomogeneous equation is

$$y(x) = ce^{-x}.$$

Then $y' = -ce^{-x}$, $y'' = ce^{-x}$ and

$$\begin{aligned} y'' - y' - 6y &= ce^{-x} + ce^{-x} - 6ce^{-x} \\ &= -4ce^{-x} \stackrel{!}{=} e^{-x}. \end{aligned}$$

The particular solution is

$$y_p(x) = -\frac{1}{4}e^{-x}$$

and the general solution is

$$y(x) = y_h(x) + y_p(x) = c_1e^{3x} + c_2e^{-2x} - \frac{1}{4}e^{-x}\,.$$

We study the solution on the interval $[0, \infty[$:

$$\begin{array}{ll} e^{\alpha x} > 0 & \text{for } \alpha \in \mathbb{R},\ x \in \mathbb{R} \\ e^0 = 1 & \\ \lim_{x\to\infty} e^{\alpha x} = \infty & \text{for } \alpha \in \mathbb{R},\ \alpha > 0 \\ \lim_{x\to\infty} e^{\alpha x} = 0 & \text{for } \alpha \in \mathbb{R},\ \alpha < 0 \end{array}$$

therefore $y(x)$ is bounded if and only if $c_1 = 0$. The condition $y(0) = 0$ yields the solution

$$y(x) = \frac{1}{4}e^{-2x} - \frac{1}{4}e^{-x}\,.$$

6 Appendix

6.1 Another chain rule for partial derivatives

We consider the path of a particle in the x, y-plane:

$$\vec{r}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

and a scalar field that depends on the time $\psi(x, y, t)$. The value of ψ in $\vec{r}(t)$ is a function $f(t)$ of t.

$$\psi(x, y, t)|_{x=x(t)} = \psi\big(x(t), y(t), t\big) =: f(t)\,.$$

The derivative $\dot{f}(t)$ is given by the chain rule for partial derivatives:

$$\begin{aligned}
\dot{f}(t) &= \frac{d}{dt}\psi\big(x(t), y(t), t\big) \\
&= \left.\frac{\partial\psi(x, y, t)}{\partial x}\right|_{\vec{r}=\vec{r}(t)} \dot{x}(t) + \left.\frac{\partial\psi(x, y, t)}{\partial y}\right|_{\vec{r}=\vec{r}(t)} \dot{y}(t) + \left.\frac{\partial\psi(x, y, t)}{\partial t}\right|_{\vec{r}=\vec{r}(t)} \\
&= \psi_x\big(x(t), y(t), t\big)\, \dot{x}(t) + \psi_y\big(x(t), y(t), t\big)\, \dot{y}(t) + \psi_t\big(x(t), y(t), t\big)\,,
\end{aligned}$$

with

$$\psi_x(x, y, t) := \frac{\partial\psi(x, y, t)}{\partial x},\ \ \psi_y(x, y, t) := \frac{\partial\psi(x, y, t)}{\partial y},\ \ \psi_t(x, y, t) := \frac{\partial\psi(x, y, t)}{\partial t}.$$

To show the rule we compute the limit

$$\dot{f}(t) = \lim_{\Delta t \to 0}\Big[\psi\big(x(t + \Delta t), y(t + \Delta t), t + \Delta t\big) - \psi\big(x(t), y(t), t\big)\Big]\,.$$

For very small Δt we have

$$\begin{aligned}
&\psi\big(x(t + \Delta t), y(t + \Delta t), t + \Delta t\big) \\
&= \psi\big(x(t) + \dot{x}(t)\Delta t,\ y(t) + \dot{y}(t)\Delta t,\ t + \Delta t\big) \\
&= \psi\big(x(t), y(t) + \dot{y}(t)\Delta t, t + \Delta t\big) + \underbrace{\psi_x\big(x(t), y(t) + \dot{y}(t)\Delta t, t + \Delta t\big)}_{X}\, \dot{x}(t)\Delta t\,.
\end{aligned}$$

Moreover

$$\begin{aligned}
&\psi\big(x(t), y(t) + \dot{y}(t)\Delta t, t + \Delta t\big) \\
&= \psi\big(x(t), y(t), t + \Delta t\big) + \underbrace{\psi_y\big(x(t), y(t), t + \Delta t\big)}_{Y}\, \dot{y}(t)\Delta t\,,
\end{aligned}$$

and

$$\psi\big(x(t), y(t), t + \Delta t\big) = \psi\big(x(t), y(t), t\big) + \underbrace{\psi_t\big(x(t), y(t), t\big)}_{T} \Delta t \,.$$

Herewith

$$\dot{f}(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Big[X\dot{x}(t)\Delta t + Y\dot{y}(t)\Delta t + T\Delta t \Big] = \lim_{\Delta t \to 0} \Big[X\dot{x}(t) + Y\dot{y}(t) + T \Big]$$

yields the assumption.
The chain rule is also written

$$\frac{d\psi}{dt} = \frac{\partial \psi}{\partial x}\dot{x} + \frac{\partial \psi}{\partial y}\dot{y} + \frac{\partial \psi}{\partial t} \,.$$

6.2 The flux through a surface

We consider a flux (of a liquid) through the surface. In the point $(x, y, z) \in \mathbb{R}^3$ the flux is given by the vector field

$$\begin{array}{rccl} \vec{b}: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x,y,z) & \longmapsto & \vec{b}(x,y,z)\,. \end{array}$$

Then the quantity of liquid flowing through the vector surface S_k during a unit time interval (if the flux doesn't change in time) can be approximated with

$$\psi_k := \langle \vec{b}, \vec{r}_u \times \vec{r}_v \rangle \, \mu(D_k)\,.$$

The flux through the surface S is

$$\psi = \int_D \langle \, \vec{b}\big(\vec{r}(u,v)\big), \big(\vec{r}_u(u,v) \times \vec{r}_v(u,v)\big) \, \rangle \, d\mu(u,v)\,.$$

This is written as

$$\int_S \langle \vec{b}, \vec{n} \rangle \, d\omega\,,$$

where

$$\vec{n} := \frac{\vec{r}_u \times \vec{r}_v}{|\vec{r}_u \times \vec{r}_v|}\,.$$

is the normal vector to S in $\vec{r}(u,v)$ and

$$\big(\vec{r}_u(u,v) \times \vec{r}_v(u,v)\big) \, d\mu(u,v) = \vec{n} \, d\omega$$

the vector *surface element*.

Example. We compute the flux of the Coulomb field

$$\vec{K}(\vec{r}) := \frac{C}{r^2} \frac{\vec{r}}{r}\,, \quad (\vec{r} \neq 0,\ r = |\vec{r}|)$$

from the inner to the outer of the sphere S_R^2. The normal vector at the point $\vec{r}$ of the sphere is

$$\vec{n} = \frac{\vec{r}}{R}$$

and herewith

$$\langle \vec{K}, \vec{n} \rangle = \langle \frac{C}{R^2} \frac{\vec{r}}{R}, \frac{\vec{r}}{R} \rangle = \frac{C}{R^2}\,.$$

Herewith

$$\begin{aligned} \psi &= \int_{S_R^2} \langle \vec{K}, \vec{n} \rangle \, d\omega = \frac{C}{R^2} \int_{S_R^2} d\omega \\ &= \frac{C}{R^2} \, \omega(S_R^2) = 4\pi \, C\,. \end{aligned}$$

The flux does not depend on R.

6.3 The orthogonal group

6.3.1 Linear groups

We already know the addition and mulitplication of matrices. We know that there is a $n \times n$-matrix Id_n such that for any $n \times n$-matrix A

$$\mathrm{Id}_n\, A = A\, \mathrm{Id}_n = A\,.$$

For which $n \times n$-matrices A does exist a $n \times n$-matrix B such that

$$A\, B = B\, A = \mathrm{Id}_n\,?$$

This matrix is called the inverse of A and denoted by A^{-1}. The inverse exists if and only if $\det A \neq 0$.

Definition 6.1. We call $\mathrm{Mat}(m \times n, \mathbb{R})$ the set (ring) of $m \times n$-matrices over $\mathbb{R}$. The *general linear group* $\mathrm{GL}(n, \mathbb{R})$ is defined to be the set (group) of invertible matrices

$$\mathrm{GL}(n, \mathbb{R}) := \left\{ A \in \mathrm{Mat}(n \times n, \mathbb{R}) \,\middle|\, \det A \neq 0 \right\}.$$

The group $\mathrm{GL}(n, \mathbb{R})$ describes the set of invertible linear mappings

$$\mathbb{R}^n \longrightarrow \mathbb{R}^n\,.$$

6.3.2 Isometries

Another group of linear applications are built by the isometries. These are mappings that preserve distances and angles. The scalar product measures angles and distances. Indeed

$$\langle x, y \rangle = ||x||\, ||y||\, \cos \alpha\,,$$

where $||x|| := \sqrt{\langle x, x \rangle}$, resp. $||y|| := \sqrt{\langle y, y \rangle}$, denotes the length of x, resp. y, and α is the angle between x and y.

Definition 6.2. A mapping

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

is an *isomety* if and only if for all $x, y \in \mathbb{R}^n$

$$\langle x, y \rangle = \langle f(x), f(y) \rangle\,.$$

Then in particular $||x|| = ||f(x)||$, $||y|| = ||f(y)||$, and the angle between x and y equals the one between $f(x)$ and $f(y)$.
Let $\tau_1, \dots, \tau_n$ be a basis of $\mathbb{R}^n$.

Definition 6.3. The basis $\tau_1, \dots, \tau_n$ of $\mathbb{R}^n$ is called an *orthogonal* basis, if for all $i, j = 1, \dots, n$

$$\langle \tau_i, \tau_j \rangle = 0 \text{ if and only if } i \neq j$$

and it is called an *orthonormal* basis, if for all $i, j = 1, \dots, n$

$$\langle \tau_i, \tau_j \rangle = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \,. \end{cases}$$

The standard basis $e_1, \dots, e_n$ is an orthonormal basis of $\mathbb{R}^n$. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be an isometry. Then for all $x, y \in \mathbb{R}^n$

$$\langle x, y \rangle = \langle f(x), f(y) \rangle$$

and in particular

$$\langle e_i, e_j \rangle = \langle f(e_i), f(e_j) \rangle \,, \qquad i, j = 1, \dots, n \,.$$

Therefore the orthonormal standard basis is mapped to an orthonormal basis $f(e_1), \dots, f(e_n)$. Let A denote the matrix corresponding to the mapping f:

$$\begin{array}{rccl} f : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & x & \longmapsto & f(x) := Ax \end{array}$$

then the column-vectors of A are the images of the standard basis under f. Therefore

$$A = (f(e_1)\,, \; \dots \;, \; f(e_n)) = (\tau_1\,, \; \dots \;, \; \tau_n)$$

and $\tau_1, \dots, \tau_n$ is an orthogonal basis of $\mathbb{R}^n$. We compute

$$A^{\mathrm{t}} A = \begin{pmatrix} \tau_1^{\mathrm{t}} \\ \vdots \\ \tau_n^{\mathrm{t}} \end{pmatrix} \begin{pmatrix} \tau_1\,, & \dots\,, & \tau_n \end{pmatrix} = \left(\langle \tau_i, \tau_j \rangle \right)_{i,j=1,\dots,n} \,.$$

If f is an isometry, then $A^{\mathrm{t}} A = \mathrm{Id}_n$.

Definition 6.4. A matrix $A \in \mathrm{GL}n, \mathbb{R}$ is *orthogonal*, if and only if it satisfies

$$A^{\mathrm{t}} A = \mathrm{Id}_n \,.$$

The group

$$\mathrm{O}(n) := \left\{A \in \mathrm{GL}(n,\mathbb{R}) | A^{\mathrm{t}}\, A = \mathrm{Id}_n\right\}$$

is called the *orthogonal* group. The group

$$\mathrm{SO}(n) := \left\{A \in \mathrm{O}(n) | \det A = 1\right\}$$

is called the *special orthogonal* group.

Remark 6.5. For $A \in \mathrm{O}(n)$ holds $\det A = \pm 1$. Let v be an eigenvector to the eigenvalue λ of the orthogonal matrix A. Then

$$\langle v, v\rangle = \langle Av, Av\rangle = \langle \lambda v, \lambda v\rangle = |\lambda|^2 \langle v, v\rangle$$

Here λ is a complex number and $|\lambda|^2 = \overline{\lambda}\,\lambda = 1$.

6.3.3 Reflections and rotations

The orthogonal group is the group of reflections and rotations in $\mathbb{R}^n$ and the special orthogonal group is the group of rotations in $\mathbb{R}^n$.
Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a rotation with angle α. Let

$$e_1 = \begin{pmatrix}1\\0\end{pmatrix}, \quad e_2 = \begin{pmatrix}0\\1\end{pmatrix}$$

be the standard basis. Then

$$f(e_1) = \begin{pmatrix}\cos\alpha\\ \sin\alpha\end{pmatrix}, \quad f(e_2) = \begin{pmatrix}-\sin\alpha\\ \cos\alpha\end{pmatrix}$$

and $f(x) = Ax$ with

$$A = \begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix} .$$

The matrix A satisfies

$$\begin{aligned} A^{\mathrm{t}}\, A &= \begin{pmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix} \\ &= \begin{pmatrix}\cos^2\alpha + \sin^2\alpha & -\cos\alpha\sin\alpha + \sin\alpha\cos\alpha\\ -\sin\alpha\cos\alpha + \cos\alpha\sin\alpha & \cos^2\alpha + \sin^2\alpha\end{pmatrix} \\ &= \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\end{aligned}$$

and

$$\det A = \det\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix} = \cos^2\alpha + \sin^2\alpha = 1 .$$

This shows that $A \in \mathrm{SO}(n)$. The eigenvalue of A are the zeros of the characteristic polynomial

$$\begin{aligned}\det(A - \lambda \mathrm{Id}_2) &= \det \begin{pmatrix} \cos\alpha - \lambda & -\sin\alpha \\ \sin\alpha & \cos\alpha - \lambda \end{pmatrix} \\ &= (\cos\alpha - \lambda)^2 + \sin^2\alpha \\ &= \lambda^2 - 2\lambda\cos\alpha + 1 \\ &= \big(\lambda - (\cos\alpha + i\sin\alpha)\big)\big(\lambda - (\cos\alpha - i\sin\alpha)\big)\,.\end{aligned}$$

Therefore the eigenvalues of A are

$$e^{i\alpha} = \cos\alpha + i\sin\alpha\,, \quad e^{-i\alpha} = \cos\alpha - i\sin\alpha\,.$$

These are complex conjugate.
We consider the reflection on the line $\{r(\cos\frac{\alpha}{2}, \sin\frac{\alpha}{2}) \mid r \in \mathbb{R}\}$. It is given by the matrix

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}\,.$$

The characteristic polynomial is

$$\begin{aligned}\det A &= \det \begin{pmatrix} \cos\alpha - \lambda & \sin\alpha \\ \sin\alpha & -\cos\alpha - \lambda \end{pmatrix} \\ &= -(\cos\alpha - \lambda)(\cos\alpha + \lambda) - \sin^2\alpha \\ &= -\cos^2\alpha + \lambda^2 - \sin^2\alpha \\ &= \lambda^2 - 1 \\ &= (\lambda - 1)(\lambda + 1)\,.\end{aligned}$$

The eigenvalue are $\lambda_1 = 1$ and $\lambda_2 = -1$. The corresponding eigenvectors are

$$v_1 = \begin{pmatrix} v_{11} \\ v_{21} \end{pmatrix}\,, \qquad v_2 = \begin{pmatrix} v_{12} \\ v_{22} \end{pmatrix}$$

with

$$\begin{pmatrix} \cos\alpha - \lambda_i & \sin\alpha \\ \sin\alpha & -\cos\alpha - \lambda_i \end{pmatrix} \begin{pmatrix} v_{1i} \\ v_{2i} \end{pmatrix} = 0\,,$$

$i = 1, 2$. Solving the equations we get

$$\begin{aligned}\begin{pmatrix} \cos\alpha - 1 & \sin\alpha \\ \sin\alpha & -\cos\alpha - 1 \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{21} \end{pmatrix} &= 0 \\ \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \begin{pmatrix} v_{11} \\ v_{21} \end{pmatrix} &= \begin{pmatrix} v_{11} \\ v_{21} \end{pmatrix} \\ \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix} &= \begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix}\end{aligned}$$

Since

$$\cos\frac{\alpha}{2} = \cos\alpha\,\cos\frac{\alpha}{2} + \sin\alpha\,\sin\frac{\alpha}{2} = \cos\bigl(\alpha - \frac{\alpha}{2}\bigr)\,,$$
$$\sin\frac{\alpha}{2} = \sin\alpha\,\cos\frac{\alpha}{2} - \cos\alpha\,\sin\frac{\alpha}{2} = \sin\bigl(\alpha - \frac{\alpha}{2}\bigr)\,.$$

For the eigenvalue -1 we have

$$\begin{pmatrix} \cos\alpha + 1 & \sin\alpha \\ \sin\alpha & -\cos\alpha + 1 \end{pmatrix} \begin{pmatrix} v_{12} \\ v_{22} \end{pmatrix} = 0$$
$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \begin{pmatrix} v_{12} \\ v_{22} \end{pmatrix} = \begin{pmatrix} -v_{12} \\ -v_{22} \end{pmatrix}$$
$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} \begin{pmatrix} -\sin\frac{\alpha}{2} \\ \cos\frac{\alpha}{2} \end{pmatrix} = \begin{pmatrix} \sin\frac{\alpha}{2} \\ -\cos\frac{\alpha}{2} \end{pmatrix}$$

Herewith the eigenvector v_1 to the eigenvalue $\lambda_1 = 1$ is

$$v_1 = \begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix}$$

and the eigenvector v_{-1} to the eigenvalue $\lambda_{-1} = -1$ is

$$v_{-1} = \begin{pmatrix} -\sin\frac{\alpha}{2} \\ \cos\frac{\alpha}{2} \end{pmatrix}\,.$$

These vectors form an orthonormal basis of $\mathbb{R}^2$, i.e. they are linearly independent and

$$||v_1|| = 1\,,\ ||v_{-1}|| = 1\,,\quad \langle v_1, v_{-1}\rangle = 0\,.$$

In the basis of eigenvectors the reflection is given by

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,.$$

In 2 dimensions the rotation is around the origin of the plane. In 3 dimensions the rotation is around an axis that goes through the origin (one-dimensional subspace). The axis is fixed by the rotation and therefore spanned by an eigenvector v to the eigenvalue 1. The axis is perpendicular to a plane E. Restricted to this plane, the rotation has the properties we described above. So the matrix with respect to the basis given by v and a basis of the plane is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}\,.$$

The matrix of the reflection on the plane E, i.e. that fixes E and maps $v \mapsto -v$ is

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

If we combine both mappings (where in this case we get the same result if we reflect or if we rotate first).

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} .$$

There is a theorem that for every orthogonal matrix there is a basis in which it looks like a matrix with blocks of

$$1, \quad -1, \quad \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \quad \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} .$$

on the diagonal.